

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:31:29 ; Search time 140.273 Seconds
(without alignments)
408.065 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPKHSHQHPGHHPAH.....PSFPLPHKHPKPNQPPF 148

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:.*
1: geneseqp1980s:.*
2: geneseqp1990s:.*
3: geneseqp2000s:.*
4: geneseqp2001s:.*
5: geneseqp2002s:.*
6: geneseqp2003as:.*
7: geneseqp2003bs:.*
8: geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	966	100.0	148	5	ABB79807 Rabbit hi
2	966	100.0	148	5	Abb79806 Human his
3	966	100.0	148	8	ADH10411 Human HPR
4	966	100.0	525	5	Abb79804 Human his
5	966	100.0	525	8	ADH10409 Human HPR
6	949	98.2	525	8	ADH76897 Human pro
7	949	98.2	525	8	ADL12335 Human ste
8	543	56.2	526	5	Abb79805 Rabbit hi
9	543	56.2	526	8	ADH10410 Rabbit HP
10	302	31.3	75	5	Abb07123 HRGP thro
11	280.5	29.0	101	8	ADH10412 Rabbit HP
12	254.5	26.3	245	4	Abb28019 Novel hum
13	251.5	26.0	330	4	Abb25331 Novel hum
14	248.5	25.7	406	4	Abb25331 Novel hum
15	242.5	25.1	233	4	Abb19767 Novel hum
16	242.5	25.1	378	8	ABO60258 Human gen
17	241.5	25.0	351	2	AAR24393 Sequence
18	240	24.8	491	4	ABG09949 Novel hum
19	230	23.8	183	4	ABG25337 Novel hum
20	225.5	23.3	598	4	ABG23595 Novel hum
21	221	22.9	292	4	AAM18326 Peptide #
22	221	22.9	292	4	ABB37360 Peptide #
23	221	22.9	292	4	AAM30814 Peptide #
24	221	22.9	292	4	ABB32108 Peptide #
25	221	22.9	292	4	ABB22646 Protein #

26	221	22.9	292	4	AAM70490 Human bon
27	221	22.9	292	4	AAM58050 Human bra
28	221	22.9	292	4	ABG52171 Human liv
29	221	22.9	292	4	AAM05933 Peptide #
30	221	22.9	292	5	ABG40129 Human pep
31	219.5	22.7	227	4	ABG14399 Novel hum
32	215	22.3	124	4	ABG12879 Novel hum
33	210	21.7	82	4	AAM13580 Peptide #
34	210	21.7	82	4	ABB32509 Peptide #
35	210	21.7	82	4	AAM25978 Peptide #
36	210	21.7	82	4	ABB27363 Human pep
37	210	21.7	82	4	ABB18015 Protein #
38	210	21.7	82	4	AAM65721 Human bon
39	210	21.7	82	4	AAM53343 Human bra
40	210	21.7	82	4	ABG47362 Human liv
41	210	21.7	82	4	AAM01333 Peptide #
42	210	21.7	82	5	ABG35350 Human pep
43	209.5	21.7	324	4	ABG20509 Novel hum
44	209	21.6	304	4	ABG08412 Novel hum
45	208	21.5	86	5	ADK36683 Novel hum

ALIGNMENTS

RESULT 1
ABB79807
ID ABB79807 standard; protein; 148 AA.
XX
AC ABB79807;
XX
DT 25-NOV-2002 (first entry)
XX
DE Rabbit histidine proline rich glycoprotein His/Pro-rich domain.
XX
KW Histidine proline rich glycoprotein; HPGP; rabbit; antiangiogenic;
KW cytosolic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antithrombotic; antitumor; osteoprotectin; antitumor;
KW ophthalmological; nontoxic; neuroprotective; antiparkinsonian; chelator.
XX
OS Oryctolagus cuniculus.
XX
PN WO200264621-A2.
XX
PD 22-AUG-2002.
XX
PF 14-FEB-2002; 2002WO-US004336.
XX
PR 14-FEB-2001; 2001US-0268370P.
XX
(ATTE-) ATTENUON LLC.
XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX
DR WPI, 2002-666989/71.
DR P-PSDB; ABB84910.
XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
diagnosing or treating diseases associated with undesired cell migration,
invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX
PS Claim 1; Page 13; 82pp; English.
XX
CC The present sequence is the protein sequence of the histidine-proline-
rich (H/P) domain of rabbit anti-angiogenic histidine proline rich
glycoprotein (HPRG, see also ABB79805). Claimed anti-angiogenic
polypeptides or peptides comprise: the H/P domain (see ABB79806) of human
HPRG; an H/P domain of rabbit HPRG; a variant of these that is capable of
inhibiting angiogenesis, endothelial cell proliferation or endothelial
tube formation in vitro or in vivo; or a pentapeptide having the generic
sequence given in ABB79808. Also claimed are: chemically synthesised or
recombinantly produced peptide multimers; a diagnostically or
therapeutically labeled anti-angiogenic polypeptide, peptide or peptide

CC multimer; a diagnostically useful HPRG-related composition, comprising
 CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
 CC carrier; an antibody specific for an epitope of HPRG that is present in
 CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
 CC of the domains in a way which inhibits the anti-angiogenic activity of
 CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
 CC method for inhibiting cell migration, cell invasion, cell proliferation
 CC or angiogenesis, or for inducing apoptosis; a method for treating a
 CC subject having a disease or condition associated with undesired cell
 CC migration, invasion, proliferation, or angiogenesis; a method of
 CC stimulating or inhibiting angiogenesis in a subject; a method of
 CC detecting the presence of HPRG or its cleavage product or its peptide in
 CC a biological sample; isolated nucleic acids encoding the polypeptide,
 CC peptide or peptide multimer; an expression vector; transformed or
 CC transfected cells; a method of providing to a cell, tissue or organ an
 CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
 CC pentapeptide, or the peptide multimer; an affinity ligand useful for
 CC binding to, or isolating, an HPRG-binding molecule or cells expressing
 CC the binding molecule, comprising the polypeptide, peptide or peptide
 CC multimer, immobilised to a solid support or carrier; and a method of
 CC isolating HPRG-binding molecule, or isolating or enriching cells
 CC expressing HPRG-binding site or receptor from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays

XX Sequence 148 AA;

Query Match 100.0%; Score 966; DB 5; Length 148;
 Best Local Similarity 100.0%; Pred. No. 5.2e-80;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPKHSHEQHPGHHPHHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60
 Db 1 HPKHSHEQHPGHHPHHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60

QY 61 DFQDYGCDPPPHNQCHGCHGPPPGHLRRRGKGRPFRCRQIGSVYRPLPKGEV 120
 Db 61 DFQDYGCDPPPHNQCHGCHGPPPGHLRRRGKGRPFRCRQIGSVYRPLPKGEV 120

QY 121 LPLPEANFPSPFLPHHKHPLKPDNQPPF 148
 Db 121 LPLPEANFPSPFLPHHKHPLKPDNQPPF 148

RESULT 2
 ABB79806
 ID ABB79806 standard; protein; 148 AA.

XX ABB79806;
 XX ABB79806;
 XX 25-NOV-2002 (first entry)
 XX Human histidine proline rich glycoprotein His/Pro-rich domain.
 XX Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
 KW cystostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
 KW synecological; antiarthritic; antitumor; osteopathic; antitumor;
 KW ophthalmological; neurotropic; neuroprotective; antiparkinsonian.
 XX Homo sapiens.
 XX WO200264621-A2.
 XX 22-AUG-2002.
 XX

PF 14-FEB-2002; 2002WO-US0004336.
 PR 14-FEB-2001; 2001US-0268370P.
 XX (ATE-) ATTENUNO LLC.
 PA Donate F, Harris S, Plunkett ML, Mazar AP;
 PI WPI; 2002-666989/71.
 DR P-PSDB; ABB84910.
 XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
 PT diagnosing or treating diseases associated with undesired cell migration,
 PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
 PS Claim 1; Page 13; 82pp; English.

XX The present sequence is the protein sequence of the histidine-proline-
 CC rich (H/P) domain of human anti-angiogenic histidine proline rich
 CC glycoprotein (HPRG, see also ABB79804). Claimed anti-angiogenic
 CC polypeptides or peptides comprise: the H/P domain of human HPRG; an H/P
 CC domain of rabbit HPRG (see ABB12345); a variant of these that is capable
 CC of inhibiting angiogenesis, endothelial cell proliferation or endothelial
 CC tube formation in vitro or in vivo; or a pentapeptide having the generic
 CC sequence given in ABB79808. Also claimed are: chemically synthesised or
 CC recombinantly produced peptide multimers; a diagnostically or
 CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
 CC multimer; a diagnostically useful HPRG-related composition, comprising
 CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
 CC carrier; an antibody specific for an epitope of HPRG that is present in
 CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
 CC of the domains in a way which inhibits the anti-angiogenic activity of
 CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
 CC method for inhibiting cell migration, cell invasion, cell proliferation
 CC or angiogenesis, or for inducing apoptosis; a method for treating a
 CC subject having a disease or condition associated with undesired cell
 CC migration, invasion, proliferation, or angiogenesis; a method of
 CC stimulating or inhibiting angiogenesis in a subject; a method of
 CC detecting the presence of HPRG or its cleavage product or its peptide in
 CC a biological sample; isolated nucleic acids encoding the polypeptide,
 CC peptide or peptide multimer; an expression vector; transformed or
 CC transfected cells; a method of providing to a cell, tissue or organ an
 CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
 CC pentapeptide, or the peptide multimer; an affinity ligand useful for
 CC binding to, or isolating, an HPRG-binding molecule or cells expressing
 CC the binding molecule, comprising the polypeptide, peptide or peptide
 CC multimer, immobilised to a solid support or carrier; and a method of
 CC isolating HPRG-binding molecule, or isolating or enriching cells
 CC expressing HPRG-binding site or receptor, from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays

XX Sequence 148 AA;

Query Match 100.0%; Score 966; DB 5; Length 148;
 Best Local Similarity 100.0%; Pred. No. 5.2e-80;
 Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPKHSHEQHPGHHPHHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60
 Db 1 HPKHSHEQHPGHHPHHAHPHEHDTHQHPGHHPHGHHPHGHHPHGHHPHCH 60

QY 61 DFQDYGCDPPPHNQCHGCHGPPPGHLRRRGKGRPFRCRQIGSVYRPLPKGEV 120
 Db 61 DFQDYGCDPPPHNQCHGCHGPPPGHLRRRGKGRPFRCRQIGSVYRPLPKGEV 120

Sequence 148 AA;

angiogenic polypeptide, peptide or peptide multimer; a diagnostically useful HPRG-related composition, comprising the diagnostically labeled polypeptide, peptide or peptide multimer and a carrier; an antibody specific for an epitope of HPRG that is present in the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any of the domains in a way which inhibits the anti-angiogenic activity of HPRG or the domain, or an antigen-binding fragment of the antibody; a method for inhibiting cell migration, cell invasion, cell proliferation or angiogenesis, or for inducing apoptosis; a method for treating a subject having a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis; a method of stimulating or inhibiting angiogenesis in a subject; a method of detecting the presence of HPRG or its cleavage product or its peptide in a biological sample; isolated nucleic acids encoding the polypeptide, peptide or peptide multimer; an expression vector; transformed or transfected cells; a method of providing to a cell, tissue or organ an angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer; an affinity ligand useful for binding to, or isolating, an HPRG-binding molecule or cells expressing the binding molecule, comprising the polypeptide, peptide or peptide multimer, immobilised to a solid support or carrier; and a method of isolating HPRG-binding molecule, or isolating or enriching cells expressing HPRG-binding site or receptor, from a complex mixture. The compositions and methods are useful in diagnosing or treating a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or fractures. HPRG is especially useful in inhibiting the growth of primary tumours or metastases, and may also be used in treating neurodegenerative diseases like Alzheimer's disease, Parkinson's disease, and amyotrophic lateral sclerosis. The antibodies are stimulators of angiogenesis and are useful for promoting neovascularization in pertinent disease states, and in various immunoassays

Sequence 525 AA;

Query Match 100.0%; Score 966; DB 5; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.9e-79;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPKHHSHEQHPGHHPHAAHPHEHDTTHROHPGHHPHGHHPHGHHPHGHHPHCH 60
DB 350 HPKHHSHEQHPGHHPHAAHPHEHDTTHROHPGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGCDPPPHNQHCCHGCHGPPPGHLRRRGKGGKPPFRCRQIGSVYRLPPLRKGEV 120
DB 410 DFQDYGCDPPPHNQHCCHGCHGPPPGHLRRRGKGGKPPFRCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFSPPLPHHKHPLKPDNQPPF 148
DB 470 LPLPEANFSPPLPHHKHPLKPDNQPPF 497

RESULT 5
ADH10409
ID ADH10409 standard; protein; 525 AA.
XX AC ADH10409;
XX DT 11-MAR-2004 (first entry)
XX DE Human HPRG protein sequence.
XX KW Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritis;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnary;
KW angiogenesis; gene therapy; human.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Domain 276. .321

FT Domain /note = pro-rich domain
FT 350. .497
XX /note = His-Pro rich domain
PN WO2003077872-A2.
XX 25-SEP-2003.
XX 17-MAR-2003; 2003WO-US008060.
XX 15-MAR-2002; 2002US-0364047P.
XX (ATTE-) ATTENUON LLC.
XX McCrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for inhibiting endothelial cell migration, invasion, proliferation or angiogenesis, inducing endothelial cell apoptosis, or treating tumors or cancer.

Claim 6; SEQ ID NO 21; 117pp; English.

The invention relates to an isolated tropomyosin (Tpm)-related antiangiogenic receptor polypeptide or peptide, which is a fragment of a full-length native Tpm protein expressed on the surface of endothelial cells, or a variant of the fragment. It has a molecular mass of about 17 kDa and corresponds in its sequence to, or is a variant of, an internal fragment of a native Tpm isoform which is a binding site for antiangiogenic polypeptide agents; The isolated antiangiogenic receptor polypeptide, peptide or variant has substantially the same biochemical activity of binding to the antiangiogenic polypeptide agents, as does the native Tpm internal fragment. The antiangiogenic polypeptide agent that binds to the isolated polypeptide or peptide is human histidine-proline rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic homologue, variant, domain or fragment of human or rabbit HPRG, two chain human kininogen (HK), the D5 domain of HK, or a Tpm-binding, antiangiogenic homologue, variant, domain or fragment of the HK or its D5 domain. The Tpm-related antiangiogenic receptor polypeptide or peptide, antibodies and compositions are useful for inhibiting endothelial cell migration, invasion, proliferation or angiogenesis, for inducing endothelial cell apoptosis, or for treating tumours or cancer, diabetic retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis, psoriasis, or scleroderma. The antibody may be also used for detecting the presence of a Tpm polypeptide or peptide in a biological sample, for promoting wound healing, or for treating diseases or conditions in which increased angiogenesis is desired, e.g. coronary artery disease or peripheral artery disease. The present sequence represents a human HPRG protein sequence.

Sequence 525 AA;

Query Match 100.0%; Score 966; DB 8; Length 525;
Best Local Similarity 100.0%; Pred. No. 1.9e-79;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPKHHSHEQHPGHHPHAAHPHEHDTTHROHPGHHPHGHHPHGHHPHGHHPHCH 60
DB 350 HPKHHSHEQHPGHHPHAAHPHEHDTTHROHPGHHPHGHHPHGHHPHGHHPHCH 409

QY 61 DFQDYGCDPPPHNQHCCHGCHGPPPGHLRRRGKGGKPPFRCRQIGSVYRLPPLRKGEV 120
DB 410 DFQDYGCDPPPHNQHCCHGCHGPPPGHLRRRGKGGKPPFRCRQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFSPPLPHHKHPLKPDNQPPF 148
DB 470 LPLPEANFSPPLPHHKHPLKPDNQPPF 497

RESULT 6
ADE76897

ID ADE76897 standard; protein; 525 AA.
AC ADE76897;
DT 29-JAN-2004 (first entry)
XX Human protein expressed in a liver disorder #18.
DE human; liver disorder; hyperlipidaemia; hypertension; type II diabetes;
KW tumour; liver; inflammatory disorder; immune response disorder;
XX high-throughput screening; differential gene expression; gene therapy.
KW Homo sapiens.
OS Homo sapiens.
XX US2003108871-A1.
PN 12-JUN-2003.
XX 30-JUL-2001; 2001US-00919039.
PD 28-JUL-2000; 2000US-0222113P.
XX (KASE/) KASER M R.
PA Kaser MR;
PI WPI; 2004-031227/03.
XX N-PSDB; ADE76896.
DR Composition comprising several cDNAs that are differentially expressed in
PT treated human C3A liver cell cultures, useful for treating liver
PT disorders.
XX Claim 1; SEQ ID NO 62; 41pp; English.
PS The invention relates to a composition comprising several cDNAs that are
CC differentially expressed in a liver disorder. The composition is useful
CC for treating liver disorder such as hyperlipidaemia, hypertension, type
CC II diabetes, tumours of the liver and disorders of the inflammatory and
CC immune response. The composition is useful for a high-throughput method
CC of screening several molecules or compounds to identify a ligand which
CC specifically binds a cDNA. A protein encoded by the cDNA is useful for a
CC high-throughput method for using a protein to screen several molecules or
CC compounds to identify at least one ligand which specifically binds the
CC protein which involves combining the protein encoded by the cDNA with
CC several of molecules or compounds under conditions to allow specific
CC binding, and detecting specific binding between the protein and a
CC molecule or compound, therefore identifying a ligand which specifically
CC binds the protein. The composition is useful for detecting and
CC quantifying differential gene expression, can be used in gene therapy, to
CC formulate prognosis and to design a treatment regimen and to monitor the
CC efficacy of treatment. The present sequence represents the amino acid
CC sequence of a protein encoded by a cDNA differentially expressed in a
CC liver disorder.
XX
SQ Sequence 525 AA;
Query Match 98.2%; Score 949; DB 8; Length 525;
Best Local Similarity 98.6%; Pred. No. 6.6e-78;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 60
350 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 409
DB 61 DFQDYGCPDPHPNQHCHGCHGPPGCHLRRCGKGPFFHCRQIGSVYRLLPKRKEV 120
410 DFQDYGCPDPHPNQHCHGCHGPPGCHLRRCGKGPFFHCRQIGSVYRLLPKRKEV 469
QY 121 LPLPEANFSPFLPHKHKLKPDNQPPF 148
470 LPLPEANFSPFLPHKHKLKPDNQPPF 497
DB

RESULT 7
ADL12335
ID ADL12335 standard; protein; 525 AA.
XX
AC ADL12335;
XX
DT 06-MAY-2004 (first entry)
XX
DE Human steroid-induced C3A liver cell protein #11.
XX
KW Hepatotropic; Gene therapy; Wilson disease; liver disorder;
KW steroid therapy; cirrhosis; hepatitis; human; C3A liver cell.
XX
OS Homo sapiens.
XX
PN US6673549-B1.
XX
PD 06-JAN-2004.
XX
PF 12-OCT-2001; 2001US-00976594.
XX
PR 12-OCT-2000; 2000US-0240409P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Furness LM, Buchbinder JL;
XX
DR WPI; 2004-068610/07.
XX
XX Combination useful for preparing a composition for treating liver
PT disorders associated with steroid therapy, e.g., cirrhosis or hepatitis,
PT comprises cDNAs that are differentially expressed in response to steroid
PT treatment.
XX
PS Disclosure; SEQ ID NO 64; 141pp; English.
XX
XX The invention relates to a combination comprising cDNAs that are
CC differentially expressed in response to steroid treatment. Also included
CC are the following: a high throughput method for using a cDNA to detect
CC differential expression of nucleic acids in a sample; and a high
CC throughput method of screening molecules or compounds to identify a
CC ligand that specifically binds a cDNA. The sample is from a subject with
CC Wilson disease and comparison of a standard defines a stage of that
CC disease. The high throughput method of screening molecules or compounds
CC to identify a ligand that specifically binds a cDNA comprises: combining
CC the combination with molecules or compounds under conditions to allow
CC specific binding; and detecting specific binding between each cDNA and at
CC least one molecule or compound. The molecules or compounds are regulatory
CC proteins. The combination is useful for preparing a composition for
CC treating liver disorders associated with steroid therapy, e.g., cirrhosis
CC or hepatitis. The present sequence represents a human protein which is
CC differentially expressed in steroid-induced C3A liver cells. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format directly from USPTO
CC at seqdata.uspto.gov/sequence.html.
XX
SQ Sequence 525 AA;
Query Match 98.2%; Score 949; DB 8; Length 525;
Best Local Similarity 98.6%; Pred. No. 6.6e-78;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 60
350 HPHKHSHEQHPGHHPHAAHPHEHDTHQHPGHHPGHHPGHHPGHHPGHHPGHHPCH 409
DB 61 DFQDYGCPDPHPNQHCHGCHGPPGCHLRRCGKGPFFHCRQIGSVYRLLPKRKEV 120
410 DFQDYGCPDPHPNQHCHGCHGPPGCHLRRCGKGPFFHCRQIGSVYRLLPKRKEV 469
QY 121 LPLPEANFSPFLPHKHKLKPDNQPPF 148
121 LPLPEANFSPFLPHKHKLKPDNQPPF 148

Db 470 LPLPEANFSPFLPHHKKHLKPDQPP 497

RESULT 8

ABB79805

ID ABB79805 standard; protein; 526 AA.

AC ABB79805;

XX 25-NOV-2002 (first entry)

DE Rabbit histidine proline rich glycoprotein.

XX Histidine proline rich glycoprotein; HPRG; rabbit; antiangiogenic;

KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;

KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;

KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.

XX Oryctolagus cuniculus.

OS

XX Key Location/Qualifiers

FT Domain 251..296

FT /note: "Proline-rich domain"

FT Domain 321..421

FT /note: "Histidine-proline-rich domain, region also

FT specifically claimed in Claim 1"

XX WO200264621-A2.

PN 22-AUG-2002.

XX 14-FEB-2002; 2002WO-US004336.

PF 14-FEB-2001; 2001US-0268370P.

PR (ATTE-) ATTENUON LLC.

XX Donate F, Harris S, Plunkett ML, Mazar AP;

PI WPI; 2002-666989/71.

DR P-PSDB; ABN84911.

XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,

PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX Claim 1; Page 12-13; 82pp; English.

XX The present sequence is the protein sequence of rabbit histidine proline

CC rich glycoprotein (HPRG), a proteinaceous chelator that can be used to

CC inhibit angiogenesis and treat cancer. Claimed anti-angiogenic

CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)

CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see

CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,

CC endothelial cell proliferation or endothelial tube formation in vitro or

CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.

CC Also claimed are: chemically synthesised or recombinantly produced

CC peptide multimers; a diagnostically or therapeutically labeled anti-

CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically

CC useful HPRG-related composition, comprising the diagnostically labeled

CC polypeptide, peptide or peptide multimer and a carrier; an antibody

CC specific for an epitope of HPRG that is present in the H/P domain of

CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in

CC or an antigen-binding fragment of the antibody; a method for inhibiting

CC cell migration, cell invasion, cell proliferation or angiogenesis, or for

CC inducing apoptosis; a method for treating a subject having a disease or

CC condition associated with undesired cell migration, invasion,

CC proliferation, or angiogenesis; a method of stimulating or inhibiting

CC angiogenesis in a subject; a method of detecting the presence of HPRG or

CC its cleavage product or its peptide in a biological sample; isolated

CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an

CC expression vector; transfected or transfected cells; a method of

CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of

CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;

CC an affinity ligand useful for binding to, or isolating, an HPRG-binding

CC molecule or cells expressing the binding molecule, comprising the

CC polypeptide, peptide or peptide multimer, immobilised to a solid support

CC or carrier; and a method of isolating HPRG-binding molecule, or isolating

CC or enriching cells expressing HPRG-binding site or receptor, from a

CC complex mixture. The compositions and methods are useful in diagnosing or

CC treating a disease or condition associated with undesired cell migration,

CC invasion, proliferation, or angiogenesis, such as cancer.

CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,

CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in

CC inhibiting the growth of primary tumours or metastases, and may also be

CC used in treating neurodegenerative diseases like Alzheimer's disease,

CC Parkinson's disease and amyotrophic lateral sclerosis. The antibodies are

CC stimulators of angiogenesis and are useful for promoting

CC neovascularization in pertinent disease states, and in various

CC immunoassays

XX

XX Sequence 526 AA;

Query Match 56.2%; Score 543; DB 5; Length 526;

Best Local Similarity 57.6%; Pred. No. 4.9e-41;

Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

QY 2 PHKHSHEQHPHGHHPHA-----HH-----PHEHDTTHRQHPHGHHPH 42

DB 330 PHGHHHPGPPPHGHHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 389

QY 43 HGHHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 98

DB 390 HGHPHHPGPPPHGHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 448

QY 99 RPFHCROIGSVYRLPPLKRGVLPPLPEANFSPFLPHHKKHLKPDNQPP 148

DB 449 PFFHWRIGSVYQLPPLQKGEVLPPLPEANFSPFLPHHKKHLKPDNQPP 498

RESULT 9

ADH10410

ID ADH10410 standard; protein; 526 AA.

XX ADH10410;

AC ADH10410;

XX 11-MAR-2004 (first entry)

XX Rabbit HPRG protein sequence.

XX Tpm; tropomyosin; antiangiogenic receptor;

KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

KW antipsoriatic; dermatological; cardiant; vasotrophic; vulnerary;

XX angiogenesis; gene therapy; rabbit.

OS Oryctolagus cuniculus.

XX Key Location/Qualifiers

FT Domain 251..296

FT /note = pro-rich domain

FT Domain 321..421

FT /note = His-Pro rich domain

XX WO2003077872-A2.

PD 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PR 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

PI Mccrae K, Donate F, Juarez J, Mazar AP;

XX WPI; 2004-090604/09.
XX
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
XX Claim 6; SEQ ID NO 22; 117pp; English.
XX
XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HRGP), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein sequence.
XX
XX Sequence 526 AA;
XX
XX Query Match 56.2%; Score 543; DB 8; Length 526;
XX Best Local Similarity 57.6%; Pred. No. 4.9e-41;
XX Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;
XX
XX 2 PHKHSHEQPHGHHPHA-----HH-----PHEHDTHQPHGHHPGHHP 42
XX 330 PHGHHPGHPPPHGHHPGHPPPHGHHPGHPPPHGHHPGHPPPHGHHPGHPP 389
XX 43 HGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHP 98
XX 390 HGHHPGHPPPHGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHPGHHP 448
XX 99 RPFHCRQIGSVYRLPLRKGEVLPPLPEANFPSPPLPHKHPLKPDNQPPF 148
XX 449 FPFHWRIGSVQLPPLQKGEVLPPLPEANFPQLLNHTHPLKPEIQPPF 498
XX
XX RESULT 10
XX ABB07123
XX ID ABB07123 standard; protein; 75 AA.
XX
XX AC ABB07123;
XX
XX DT 13-MAR-2002 (first entry)
XX
XX DE HRGP thrombospondin (TSP)-binding domain fragment (residues 443-517).
XX
XX KW Thrombospondin; TSP-1; histidine-rich glycoprotein; HRGP; cytostatic;
XX cardiovascular; protein therapy; gene therapy; antisense inhibition;
XX ribozyme inhibition; antibody inhibition; angiogenesis; human; cancer;
XX TSP.
XX
XX OS Homo sapiens.
XX
XX PN

PN US2001041670-A1.
XX
XX 15-NOV-2001.
XX
XX 05-DEC-2000; 2000US-00730379.
XX
XX 06-DEC-1999; 99US-0169205P.
XX
XX (SIMA/) SIMANTOV R.
XX (SILV/) SILVERSTEIN R L.
XX
XX Simantov R, Silverstein RL;
XX
XX WPI; 2002-017220/02.
XX
XX Thrombospondin-binding molecules and fragments comprising regions of the
PT TSP-1 binding domains of Histidine-Rich Glycoproteins, useful for
PT treating cancers and cardiovascular diseases.
XX
XX Example; Fig 1; 22pp; English.
XX
XX The invention provides thrombospondin (TSP)-binding molecules and
CC fragments comprising regions of the TSP-1 binding domains of Histidine-
CC Rich Glycoproteins (HRGP). The proteins are used for the regulation of
CC cellular growth and proliferation, such as in the accretion of new blood
CC vessels (angiogenesis), particularly for treatment of cardiovascular
CC disease. They may be used for the reduction of angiogenesis, including
CC growth suppression and arrest, and apoptosis in normal development, for
CC example in embryogenesis, and in a wide range of disorders and diseases,
CC including those involving tumors, malignancies, neoplastic and other
CC pathological conditions and homeostatic imbalances in the control of
CC growth and development, in particular anal cancer, bladder cancer, small
CC cell lung cancer, non-small cell lung cancer, bone cancer, brain cancer,
CC breast cancer, cervical cancer, chondrosarcoma, clear cell adenocarcinoma
CC (DES), colorectal cancer, endometrial cancer, esophageal cancer, cancers
CC of the eye, eyelid, Kaposi's sarcoma, kidney cancer, cancer of the
CC larynx, leiomyosarcoma, leukemia, liver cancer, lung cancer, lymphoma,
CC melanoma, mesothelioma, oral cancer, ovarian cancer, pancreatic cancer,
CC prostate cancer, skin cancer, squamous cell cancer, stomach cancer,
CC testicular cancer, thyroid cancer, hepatoma, neuroendocrine cancer,
CC liposarcoma, head and neck cancer and a cholangiocarcinoma. The present
CC sequence represents the TSP-binding motif of HRGP
XX
XX Sequence 75 AA;
XX
XX Query Match 31.3%; Score 302; DB 5; Length 75;
XX Best Local Similarity 100.0%; Pred. No. 5.3e-20;
XX Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 97 GPRPHCRQIGSVYRLPLRKGEVLPPLPEANFPSPPLPHKHPLKPDNQPPF 148
XX 1 GPRPHCRQIGSVYRLPLRKGEVLPPLPEANFPSPPLPHKHPLKPDNQPPF 52
XX
XX RESULT 11
XX ADH10412
XX ID ADH10412 standard; protein; 101 AA.
XX
XX AC ADH10412;
XX
XX DT 11-MAR-2004 (first entry)
XX
XX DE Rabbit HPRG protein H/P rich domain.
XX
XX KW Tpm; tropomyosin; antiangiogenic receptor;
XX histidine-proline rich glycoprotein; HRGP; cytostatic; antidiabetic;
XX ophthalmological; antiinflammatory; gynaecological; antiarthritic;
XX antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
XX angiogenesis; gene therapy; rabbit.
XX
XX OS Oryctolagus cuniculus.
XX
XX PN W02003077872-A2.

XX 25-SEP-2003.
XX 17-MAR-2003; 2003WO-US008060.
XX 15-MAR-2002; 2002US-0364047P.
XX (ATTE-) ATTENUON LLC.
XX McCreae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX Claim 6; SEQ ID NO 24; 117pp; English.
XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX Sequence 101 AA;
SQ
Query Match 29.0%; Score 280.5; DB 8; Length 101;
Best Local Similarity 51.6%; Pred. No. 6.4e-18;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;
QY 2 PHKHSHSQHPRGHPRH-----HH-----PHNDTHRQHPRGHPRGHPR 42
DB 10 PHGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPR 69
QY 43 HGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPR 73
DB 70 HGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPR 100
RESULT 12
ABG28019
ID ABG28019 standard; protein; 245 AA.
XX AC ABG28019;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #28010.
DE Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX

KW food supplement; medical imaging; diagnostic; genetic disorder.
XX Homo sapiens.
XX WO200175067-A2.
XX 11-OCT-2001.
XX 30-MAR-2001; 2001WO-US008631.
XX 31-MAR-2000; 2000US-00540217.
XX 23-AUG-2000; 2000US-00649167.
XX (HYSE-) HYSEQ INC.
XX Drmanac RT, Liu C, Tang YT;
XX WPI; 2001-639362/73.
XX N-PSDB; AAS92206.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX Claim 20; SEQ ID NO 58378; 103pp; English.
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 245 AA;
Query Match 26.3%; Score 254.5; DB 4; Length 245;
Best Local Similarity 48.2%; Pred. No. 3.6e-15;
Matches 41; Conservative 2; Mismatches 39; Indels 3; Gaps 1;
QY 1 HPHKHSHSQHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGHPRGH 57
DB 25 HYHHHHHHHHYHRRHHHHPRHHPRHHYHHTHHHHHHHHYHQRHHHHHHHHH 84
QY 58 HCHDFQDYGPCDPPRHNGHCHGH 82
DB 85 HHHHHHQQHRRHHHHHHHHHHHHHHHH 109
RESULT 13
ABG25331
ID ABG25331 standard; protein; 330 AA.
XX AC ABG25331;
XX 18-FEB-2002 (first entry)
XX Novel human diagnostic protein #25322.
DE

ABG19767
ID ABG19767 standard; protein; 233 AA.
AC
XX ABG19767;
XX
DT 18-FEB-2002 (first entry)
XX
DE Novel human diagnostic protein #19758.
XX
XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.
XX
XX Homo sapiens.
OS
XX WO200175067-A2.
PN
XX 11-OCT-2001.
PD
XX 30-MAR-2001; 2001WO-US008631.
PF
XX 31-MAR-2000; 2000US-00540217.
PR
XX 23-AUG-2000; 2000US-00649167.
XX
XX (HYSE-) HYSEQ INC.
PA
XX Drmanac RT, Liu C, Tang YT;
PI
XX WPI; 2001-639362/73.
DR N-PSDB; AAS83954.
XX
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
XX Claim 20; SEQ ID NO 50126; 103pp; English.
PS
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
CC and in recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activity. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 233 AA;

Query Match 25.1%; Score 242.5; DB 4; Length 233;
Best Local Similarity 35.0%; Pred. No. 4.2e-14;
Matches 57; Conservative 8; Mismatches 71; Indels 27; Gaps 7;
QY 1 HPHKHSHQEQ-HPHGHHPHAHHPHEH-----DTRQHPHGHHPHGH-----HPGHHPHG 49
DB 52 HPHHHHHHHNHH 111
QY 50 HHPHGHHPHCHDFQYGPDPHPHGHCHGHGPPGPHLRRRGPGKGRPRPFHCRQIGSV 109
DB 112 HYSHHHHHPHH 160

QY 110 YRLPPLRKGEV-----LPLPEANFFSFLPLHHHKLKPDNQPF 148
DB 161 -AFGALEDSAVARTGLPLPEGVLCGAPTAAASAAVPPPAVEAP 202
Search completed: June 15, 2005, 13:47:11
Job time : 142.273 secs


```

; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6727066 085596CD1
US-09-919-039-62

Query Match          98.2%; Score 949; DB 4; Length 525;
Best Local Similarity 98.6%; Pred. No. 2.1e-85;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 HPHKHHSEQHGHGHHPHAHHPHEHDTHQHPHGHHHPGHGHHHPGHGHHHPGHCH 60
   |||||||
Db 350 HPHKHHSEQHGHGHHPHAHHPHEHDTHQHPHGHHHPGHGHHHPGHGHHHPGHCH 409

QY 61 DFQDYGCPDPPHNOGHCHGHPGPPGHLRRRGPGKGRPPFHCRCQIGSVYRLPPLRKGEV 120
   |||||||
Db 410 DFQDYGCPDPPHNOGHCHGHPGPPGHLRRRGPGKGRPPFHCRCQIGSVYRLPPLRKGEV 469

QY 121 LPLPEANFFSFLPHHKHPLKPDNOFPF 148
   |||||||
Db 470 LPLPEANFFSFLPHHKHPLKPDIDQFPF 497

RESULT 3
US-08-785-636-1
; Sequence 1, Application US/08785636
; Patent No. 6027942
; GENERAL INFORMATION:
; APPLICANT: Yip, Tai-Tung
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Method and Apparatus for Desorption and Ionization of
; TITLE OF INVENTION: Analytes
; FILE REFERENCE: Hutchens
; CURRENT APPLICATION NUMBER: US/08/785,636
; CURRENT FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/068,896
; EARLIER FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: peptide

US-08-785-636-1

Query Match          19.2%; Score 185; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GHHPGHHPGHHPGHHPGHHPGHHPH 58
   |||||||
Db 1 GHHPGHHPGHHPGHHPGHHPGHHPH 25

RESULT 4
US-09-095-407-2
; Sequence 2, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; APPLICANT: Yip, Tai-Tung
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; FILE REFERENCE: Hutchens
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-2

Query Match          19.2%; Score 185; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-095-407-2

Query Match          19.2%; Score 185; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 34 GHHPGHHPGHHPGHHPGHHPGHHPH 58
   |||||||
Db 1 GHHPGHHPGHHPGHHPGHHPGHHPH 25

RESULT 5
US-09-742-494-2
; Sequence 2, Application US/09742494
; Patent No. 6528320
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-2

Query Match          19.2%; Score 185; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 9.7e-12;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 34 GHHPHGHHPHGHHPHGHHPH 58
|||
Db 1 GHHPHGHHPHGHHPHGHHPH 25

RESULT 6
 US-09-809-657-2
 ; Sequence 2, Application US/09809657
 ; Patent No. 6734022
 ; GENERAL INFORMATION:
 ; APPLICANT: HUTCHENS, WILLIAM
 ; APPLICANT: YIP, TAI-TUNG
 ; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
 ; FILE REFERENCE: P00798USF/09306611
 ; CURRENT APPLICATION NUMBER: US/09/809,657
 ; CURRENT FILING DATE: 2001-03-15
 ; PRIOR APPLICATION NUMBER: 08/068,896
 ; PRIOR FILING DATE: 1993-05-28
 ; NUMBER OF SEQ ID NOS: 2
 ; SOFTWARE: Patent in version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 26
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 US-09-809-657-2

Query Match	19.2%	Score 185;	DB 4;	Length 26;
Best Local Similarity	100.0%;	Pred. No. 9.7e-12;		
Matches 25;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 34 GHHPHGHHPHGHHPHGHHPH 58
| | | | | | | | | |
Dd 1 GHHPHGHHPHGHHPHGHHPH 25

```

RESULT 7
US-09-507-323B-13
; Sequence 13, Application US/09507323B
; Patent No. 6274345
; GENERAL INFORMATION:
; APPLICANT: Choi, Jong Hyun
; APPLICANT: Lee, Sang Yup
; APPLICANT: Xu, Zhaochui
; TITLE OF INVENTION: EXPRESSION VECTORS COMPRISING A GENE
; TITLE OF INVENTION: ENCODING OMPC DERIVED FROM E. COLI AS
; TITLE OF INVENTION: MOTIF
; FILE REFERENCE: HYLEE39.001AUS
; CURRENT APPLICATION NUMBER: US/09/507,323B
; CURRENT FILING DATE: 2000-02-18
; PRIOR APPLICATION NUMBER: KR 10-1999-0005773
; PRIOR FILING DATE: 1999-02-22
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 167
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Histidine linker.
US-09-507-323B-13

```

Query Match	19.1%	Score 184.5;	DB 3;	Length 167;
Best Local Similarity	32.7%;	Pred. No. 7.9e-11;		
Matches 48;	Conservative	2;	Mismatches 50;	Indels 47;
				Gaps 9;

1	HPRKHS-----HEQH-----PGNH-----PHANPRHND-----THRQH--	31
Qy		
8	NNNNHSLDPSGNNHHHNSGLDPSGNNHHHNSGLDPSGNNHHHNSGLDPSGNNHHHNSG	67
Db		
32	--PGNHHPGH--HPRGHHPGH--HPRGHHPCHDFOD-----YCPDPPRHNQ	75
Qy		
68	LDPSGNNHHHNSGLDPSGNNHHHNSGLDPSGNNHHHNSGLDPSGNNHHHNSGLDPSGNN	127
Db		

QY	76	GHCHGHGPPPPGHLRRRGGKGP	PFH	102
Db	128	HH--HSGLDPSGHHHHHSGLDPSGH		152

```

RESULT 8
US-09-902-540-13395
; Sequence 13395, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13395
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13395

```

Query Match	18.6%	Score 180;	DB 4;	Length 163;
Best Local Similarity	40.0%;	Pred. No. 2.1e-10;		
Matches 36;	Conservative	3;	Mismatches 19;	Gaps 6;
			Indels 32;	

Qy	6	HSEQHDPHGHPHAAHPHEHDTHRQHPGHHPHGH-HPHGHPHGHHHPHGHHHPCHDFO	64
Dh	2	HNPD-HPNHGPPGHHHPHG-----PBGHCDDHTTNNHDP-----HHHPHGHD---	48

Qy 65 YGPCDPPPHNQGHCHGHPPPGHLRRRG 94
Dp 49 -----HGPAEEH-KARAP 60

```

RESULT 9
US-09-543-681A-6151
; Sequence 6151, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID S
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543.681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 6151
; LENGTH: 400
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-6151

```

```
Query Match      17.9%; Score 172.5; DB 4; Length 400;  
Best Local Similarity   29.9%; Pred.No. 3e-09;  
Matches         41; Conservative    4; Mismatches     23; Indels       69; Gaps        9;
```

QY	1 HPHKHS--NEOHPRGCHHPPHHHPHE-HDTHRQH-----PH 33 : :
Db	28 HDHHRSHSDHDHGHGGHGHHHKKHGDHHEHNATPANTVHKYDKSEQRKHNYETH 87 :
QY	34 G-----HHPHGHC-----HPHGC---HHPHGHC-HPHCGC-HPHCH 60
Db	88 GPPIIIHHHYNSGVHLFHYDAQQAQQEAFVEHHHGHDDHHAHSHETHTSHEHESH 147
QY	61 DFQDYGPCDPppphNOGH 77

QY 5 HHSHE-----QHPHGHHPH-----AHHPHE-----HDTHQHP----- 32
Db 766 HRHEGPGGGMSSGHRPHEGPGGGMSSGHRPHEGPGGGMSSGHRPHEGPGGGMSSG 825
QY 33 HGHHPH-----GHHPH-----GHHPH-----GH-HPHGHHPHCHDFQDYG 66
Db 826 SGHRPHEGPGGGMAGGHRPHEGPGGGMSSGHRPHEGPGGGMSSGHRPHEGPGGGMSSG 877
QY 67 PCPPPHNQHCHCHG-GRPP-CHLRRRGPGKG 97
Db 878 --DVP-----GHRGHRGPPPPHHRGHDGPGHG 904

RESULT 14
US-09-248-796A-19531
; Sequence 19531, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstein et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 19531
; LENGTH: 971
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-19531

Query Match 16.2%; Score 156.5; DB 4; Length 971;
Best Local Similarity 27.8%; Pred. No. 2.9e-07;
Matches 35; Conservative 6; Mismatches 26; Indels 59; Gaps 4;

QY 2 PHKHSHEQ-----HPHGHHPHAH----- 21
Db 298 PHPDHLHQVGYPPPPPPPPPLHHHHHYPSPHHPGFGFSPPPGPGPGPPVPPK 357
QY 22 PHEHDT-----HROHPGHHPHCHHHPGHHPGHHPGHHPGHHPCHDFQ 63
Db 358 PHYQSAPPSPKFKRHPKXFKQNNKGHR-HRHHPHLHHHPPPPHHHHHHPRDFQ 416
QY 64 DYGPDC 69
Db 417 QHNGQD 422

RESULT 15
US-09-281-766-19
; Sequence 19, Application US/09281766
; Patent No. 6376196
; GENERAL INFORMATION:
; APPLICANT: Conrad, Patricia C.
; APPLICANT: Louie, Kitland
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Recombinant Neospora Antigens and Their Uses
; FILE REFERENCE: 023070-082510US
; CURRENT APPLICATION NUMBER: US/09/281,766
; CURRENT FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 08/645,951
; PRIOR FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn ver. 2.1
; SEQ ID NO 19
; LENGTH: 865
; TYPE: PRT
; ORGANISM: Neospora sp.
; OTHER INFORMATION: NC-p65 cDNA

US-09-281-766-19
Query Match 16.1%; Score 155.5; DB 3; Length 865;
Best Local Similarity 29.2%; Pred. No. 3.2e-07;
Matches 47; Conservative 9; Mismatches 88; Indels 17; Gaps 6;

QY 2 PHKHSHEQHPHGHHPHAHHPHEHDTHQHPGHHPHGHHPGHHPGHHPGHHPGHHPGHHPCHD 61
Db 571 PDRQHP 630
QY 62 FQDYG-CDPP---PHNQHCHCHGHPGPPGHLRRRGPGKGP---RPF---HCRIGSV 109
Db 631 PRPNPPGASPPSSPPNQH-PHP 689
QY 110 YELP-----PLRGEVLPLEAFNFPSPPLPHKXHPKPDNQP 146
Db 690 HPYGYNLFYFTYHQSPLFYGYGRDPCPCASHPYADDSP 730

Search completed: June 15, 2005, 13:53:03
Job time : 38.8514 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:51:11 ; Search time 124.225 Seconds
(without alignments)
456.700 Million cell updates/sec

Title: US-10-074-225A-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHPGHHPHAH.....PSFPLPHKHPKPDNQPPP 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : Published Applications AA:*
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	966	100.0	148	14 US-10-074-225A-5	Sequence 5, Appli
2	966	100.0	525	17 US-10-074-225A-1	Sequence 1, Appli
3	966	100.0	525	17 US-10-868-577A-42	Sequence 42, Appli
4	949	98.2	525	10 US-09-919-039-62	Sequence 62, Appli
5	543	56.2	526	14 US-10-074-225A-3	Sequence 3, Appli
6	325	33.6	67	15 US-10-424-599-166051	Sequence 166051,
7	302	31.3	75	9 US-09-730-379B-1	Sequence 1, Appli
8	280.5	29.0	101	14 US-10-074-225A-6	Sequence 6, Appli
9	242.5	25.1	378	14 US-10-029-386-33892	Sequence 33892, A
10	221	22.9	292	9 US-09-864-761-37944	Sequence 37944, A
11	213	22.0	38	9 US-09-730-379B-9	Sequence 9, Appli

12	213	22.0	38	9 US-09-730-379B-13	Sequence 13, Appli
13	210	21.7	82	9 US-09-864-761-33313	Sequence 33313, A
14	199.5	20.7	130	15 US-10-389-586-2423	Sequence 2423, Ap
15	199.5	20.7	133	15 US-10-389-566-2424	Sequence 2424, Ap
16	199.5	20.7	265	15 US-10-389-566-2422	Sequence 2422, Ap
17	192.5	19.9	342	15 US-10-424-599-281824	Sequence 281824, A
18	191.5	19.8	156	9 US-09-864-761-41679	Sequence 41679, A
19	185	19.2	26	9 US-09-809-657-2	Sequence 2, Appli
20	185	19.2	26	9 US-09-848-512A-2	Sequence 2, Appli
21	185	19.2	26	9 US-09-742-494-2	Sequence 4, Appli
22	185	19.2	26	14 US-10-153-312A-4	Sequence 37352, A
23	185	19.2	78	9 US-09-864-761-37352	Sequence 191894, A
24	183	18.9	96	16 US-10-425-115-191894	Sequence 2, Appli
25	183	18.9	213	17 US-10-871-841-2	Sequence 185164, A
26	183	18.9	510	16 US-10-425-115-185164	Sequence 45941, A
27	182	18.8	253	15 US-10-282-122A-45941	Sequence 33727, A
28	179.5	18.6	87	9 US-09-864-761-33727	Sequence 34744, A
29	179.5	18.6	87	9 US-09-864-761-34744	Sequence 825, App
30	178.5	18.5	140	15 US-10-389-566-825	Sequence 314599, A
31	176	18.2	636	16 US-10-425-115-314599	Sequence 69999, A
32	175	18.1	430	15 US-10-425-114-69999	Sequence 207055, A
33	172	17.8	153	16 US-10-425-115-207055	Sequence 186228, A
34	172	17.8	314	15 US-10-424-599-186228	Sequence 160290, A
35	171.5	17.8	144	15 US-10-424-599-160290	Sequence 307018, A
36	167.5	17.3	109	16 US-10-425-115-307018	Sequence 1014, App
37	166	17.2	153	14 US-10-083-357-1014	Sequence 130, App
38	164.5	17.0	695	10 US-09-890-688-130	Sequence 37882, A
39	164	17.0	409	9 US-09-864-761-37882	Sequence 31185, A
40	164	17.0	108	14 US-10-029-386-31185	Sequence 14106, A
41	164	17.0	574	14 US-10-156-761-14106	Sequence 33, Appli
42	162	16.8	476	15 US-10-360-849A-33	Sequence 320892, A
43	161	16.7	371	16 US-10-425-115-320892	Sequence 76, Appli
44	160.5	16.6	366	15 US-10-406-686A-76	Sequence 122803, A
45	158	16.4	139	16 US-10-437-963-122803	

ALIGNMENTS

RESULT 1
US-10-074-225A-5
; Sequence 5, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; FILE OF INVENTION: 38342-178463
; FILE REFERENCE: 38342-178463
; CURRENT FILING DATE: 2002-02-14
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 5
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-5

Query Match	100.0%	Score 966;	DB 14;	Length 148;
Best Local Similarity	100.0%	Pred. No. 1.1e-70;		
Matches 148;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	HPHKHSHEQHPGHHPHAHHPHEDHTRQHPGHHPHGHHPHGHHPHGHHPHCH	60	
DB	1	HPHKHSHEQHPGHHPHAHHPHEDHTRQHPGHHPHGHHPHGHHPHGHHPHCH	60	
QY	61	DPQDYGPDPHPNCGHCCHGPPGHLRRGCKGPPHCRQIGSVYRLPKRGV	120	

Db 61 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 120
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
RESULT 2
US-10-074-225A-1
; Sequence 1, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-1
Query Match 100.0%; Score 966; DB 14; Length 525;
Best Local Similarity 100.0%; Pred. No. 3.4e-70;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 60
Db 350 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 409
QY 61 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db 410 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 469
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db 470 LPLPEANFSPFLPHHKHPLKPDNQPPF 497
RESULT 3
US-10-868-577A-42
; Sequence 42, Application US/10868577A
; Publication No. US20050032697A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al.
; TITLE OF INVENTION: HEPARIN BINDING VEGFR-3 LIGANDS
; FILE REFERENCE: 28967/39359A
; CURRENT APPLICATION NUMBER: US/10/868,577A
; CURRENT FILING DATE: 2004-06-14
; PRIOR APPLICATION NUMBER: US 60/478,390
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 10/669,176
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-868-577A-42
Query Match 100.0%; Score 966; DB 17; Length 525;
Best Local Similarity 100.0%; Pred. No. 3.4e-70;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 60
Db 350 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 409
QY 61 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db 410 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 469
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db 470 LPLPEANFSPFLPHHKHPLKPDNQPPF 497
RESULT 4
US-09-919-039-62
; Sequence 62, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; FEATURE:
; OTHER INFORMATION: Incyte ID No. US20030108871A1 0855596CD1
US-09-919-039-62
Query Match 98.2%; Score 949; DB 10; Length 525;
Best Local Similarity 98.6%; Pred. No. 8.1e-69;
Matches 146; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 60
Db 350 HPHKHSHEQHPGHGHPHAAHPHEHDTHQHPGHGHPHGHHPHGHHPHGHHPHGHHPHCH 409
QY 61 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 120
Db 410 DFQDYGCDPPPHNQCHCHGHPGPPGHLRRRPGKGRPFHCRQIGSVYRLPPLRKGEV 469
QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148
Db 470 LPLPEANFSPFLPHHKHPLKPDNQPPF 497
RESULT 5
US-10-074-225A-3
; Sequence 3, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3

```
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match      56.2%; Score 543; DB 14; Length 526;
Best Local Similarity 57.6%; Pred. No. 5.7e-36;
Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

QY  2 PHKHSHSQHPGHGHPHA-----HH-----PHEHDTTHRQHPHGHHPGHHP 42
Db  330 PHGHHPHGPPPHGHPHGHGPPPHGHPHGHGPPPHGHPHGHGPPPHGHPHGHGPP 389

QY  43 HGHHPHGHHPHGHHPHGHGPDQYCPDPPHNOG----HCCGHGPPPHGHLRBRGPKGP 98
Db  390 HGHPPHGHPPHGHGHPHGHGPDQYCPDPPHNOG----HCCGHGPPPHGHLRBRGPKGP 98

QY  99 RPHCHRCQIGSVYRLPLRKGEVLPLPEANFPSPPLPHHKLPLKPDNQPPF 148
Db  449 FPHWRIGSVYQLPQKGEVLPLPEANFPQLLNRNHTHPLKPEIQPPF 498

RESULT 6
US-10-424-599-166051
; Sequence 166051, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; PRIOR FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 166051
; LENGTH: 67
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(67)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_120960C.1.bep
US-10-424-599-166051

Query Match      33.8%; Score 325; DB 15; Length 67;
Best Local Similarity 92.2%; Pred. No. 3.7e-19;
Matches 47; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY  1 HPHKHSHSQHPGHGHPHGHGPDQYCPDPPHNOG----HCCGHGPPPHGHLRBRGPKGP 51
Db  17 HPHKHSHSQHPGHGHPHGHGPDQYCPDPPHNOG----HCCGHGPPPHGHLRBRGPKGP 51

RESULT 7
US-09-730-379B-1
; Sequence 1, Application US/09730379B
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; PRIOR FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
```

```
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
US-09-730-379B-1

Query Match      31.3%; Score 302; DB 9; Length 75;
Best Local Similarity 100.0%; Pred. No. 3e-17;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY  97 GPRPFCHRCQIGSVYRLPLRKGEVLPLPEANFPSPPLPHHKLPLKPDNQPPF 148
Db  1 GPRPFCHRCQIGSVYRLPLRKGEVLPLPEANFPSPPLPHHKLPLKPDNQPPF 52

RESULT 8
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match      29.0%; Score 280.5; DB 14; Length 101;
Best Local Similarity 51.6%; Pred. No. 2.2e-15;
Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;

QY  2 PHKHSHSQHPGHGHPHA-----HH-----PHEHDTTHRQHPHGHHPGHHP 42
Db  10 PHGHHPHGPPPHGHPHGHGPPPHGHPHGHGPPPHGHPHGHGPPPHGHPHGHGPP 69

QY  43 HGHHPHGHHPHGHHPHGHGPDQYCPDPPH 73
Db  70 HGHHPHGHPPPHGHPHGHGPDQYCPDPPH 100

RESULT 9
US-10-029-386-33892
; Sequence 33892, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G
; FILE REFERENCE: AEOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; PRIOR FILING DATE: 2001-12-20
; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
```

```

; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 37944
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005414.2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 6.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 6.1
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 5.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.7
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 7.9
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 11
US-09-864-761-37944

Query Match          22.9%; Score 221; DB 9; Length 292;
Best Local Similarity 44.1%; Pred. No. 3.7e-10;
Matches 41; Conservative 3; Mismatches 31; Indels 18; Gaps 4;

QY 3 HKHSHEQHPGHHPAHHPHEHDTHQHPGHGHPHGHHPGHGH-HPHGHPHGHHPHCHD 61
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 26 HHHHHHDYDHH-HHHHHHHYHHHHYHHHPHHLNHHHHHHHHHHYHHYHHHHHHHHHH- 83

QY 62 FQDYGCDPPHPNQHCCHGHPGPPCHLRRGP 94
   |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 84 -----PHHHNHHHHH-----HRRRHY 101

RESULT 11
US-09-730-379B-9
; Sequence 9, Application US/09730379B
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CON
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
US-09-730-379B-9

Query Match          22.0%; Score 213; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GRRPHCRIGSVYRLPRLKGEVLPLPEANFPSPFLP 134
Db 1 GRRPHCRIGSVYRLPRLKGEVLPLPEANFPSPFLP 38

```

```
; Patent No. US20010041670A1
; GENERAL INFORMATION:
; APPLICANT: Simantov M.D., Ronit
; APPLICANT: Silverstein M.D., Roy L.
; TITLE OF INVENTION: THROMBOSPONDIN-BINDING REGION OF HISTIDINE-RICH
; TITLE OF INVENTION: GLYCOPROTEIN AND METHODS OF USE
; FILE REFERENCE: 955-7P/CN
; CURRENT APPLICATION NUMBER: US/09/730,379B
; CURRENT FILING DATE: 2000-12-05
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 13
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION: PUBLICATION INFORMATION
; JOURNAL: Biochemistry
; VOLUME: 25
; ISSUE: 8
; PAGES: 2220-2225
; DATE: 1986
; DATABASE ACCESSION NUMBER: GenBank/P04196
; US-09-730-379B-13

Query Match      22.0%; Score 213; DB 9; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 97 GPRPFCHRCQIGSVYRPLPKRGVLPPLPEANFSPFLP 134
      |||||
Db 1 GPRPFCHRCQIGSVYRPLPKRGVLPPLPEANFSPFLP 38

RESULT 13
; Sequence 33313, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
```

```
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 33313
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007263.3
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 8.3
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 7
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 8.1
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 7.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 9.5
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 9.2
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 7.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 6.1
; OTHER INFORMATION: SWISSPROT HIT: Q12329, EVALUAE 4.90e+00
; US-09-864-761-33313

Query Match      21.7%; Score 210; DB 9; Length 82;
Best Local Similarity 56.9%; Pred. No. 9e-10;
Matches 33; Conservative 1; Mismatches 22; Indels 2; Gaps 2;

QY 1 HPHKHSHEQHPGHGHPHAHHPHEDTHRQHPGHGHPGHGHPGHGHPGHGHPH 58
      |||||
Db 27 HWHHHHHHHYHHDDHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 82

RESULT 14
; US-10-389-566-2423
; Sequence 2423, Application US/10389566
; Publication No. US20040025202A1
; GENERAL INFORMATION:
; APPLICANT: Monsanto Technology, LLC
; APPLICANT: Laurie, Cathy C
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants
; FILE REFERENCE: 38-77(52900)D
; CURRENT APPLICATION NUMBER: US/10/389,566
; CURRENT FILING DATE: 2003-03-31
; PRIOR APPLICATION NUMBER: US 60/365,301
; PRIOR FILING DATE: 2002-03-15
; PRIOR APPLICATION NUMBER: US 60/391,786
; PRIOR FILING DATE: 2002-06-25
; PRIOR APPLICATION NUMBER: US 60/392,018
; PRIOR FILING DATE: 2002-06-26
; NUMBER OF SEQ ID NOS: 2459
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2423
; LENGTH: 130
; TYPE: PRT
; ORGANISM: Bacillus anthracis
; US-10-389-566-2423

Query Match      20.7%; Score 199.5; DB 15; Length 130;
Best Local Similarity 44.6%; Pred. No. 9.7e-09;
Matches 45; Conservative 2; Mismatches 35; Indels 19; Gaps 8;

QY 1 HPHKHSHEQHPGHGHPHAHHPHEDTHRQHPGHGHPGHGHPGHGHPGHGHPH 52
      |||||
Db 9 HCHGHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHHH 68
```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:37:20 ; Search time 28.5301 Seconds
(without alignments)
499.124 Million cell updates/sec

Title: US-10-074-225a-5
Perfect score: 966
Sequence: 1 HPHKHSHEQHGHGHHPHAH.....PSFPLPHKHPLKPDNQPPF 148

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: Pirl:.*
2: Pirl:.*
3: Pirl:.*
4: Pirl:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	966	100.0	525	1 KGHUGH	histidine-rich gly
2	503.5	52.1	445	2 A60488	histidine-rich gly
3	241.5	25.0	351	1 KGZQHL	histidine-rich gly
4	217	22.5	140	2 A54523	histidine-rich pro
5	193	20.0	735	2 T45059	hypothetical prote
6	176	18.2	1571	2 T14155	zinc finger protei
7	173.5	18.0	471	2 T33997	hypothetical prote
8	171.5	17.8	657	2 A29454	knob-associated hi
9	170.5	17.7	380	2 H83994	hypothetical prote
10	167	17.3	361	2 F87286	cation efflux fami
11	166.5	17.2	369	2 I78877	CAMP responsive el
12	166.5	17.2	508	2 A54777	CAMP responsive elem
13	166.5	17.2	515	2 T23089	hypothetical prote
14	166.5	17.2	1378	2 T30173	zinc finger protei
15	162	16.8	436	2 I49714	MHC H-2K/t-w5-link
16	160	16.6	437	2 A32372	female-specific do
17	160	16.6	549	2 B32372	male-specific doub
18	158	16.4	428	2 T27544	zinc resistance pr
19	155.5	16.1	774	2 T39539	alpha-amylase homo
20	155	16.0	199	2 T48099	hypothetical prote
21	149.5	15.5	659	2 A36664	S59/2 homeotic pro
22	149.5	15.5	1002	2 T43236	carboxypeptidase C
23	148.5	15.4	242	2 H82061	hypothetical prote
24	147	15.2	373	2 AD0262	probable membrane
25	147	15.2	375	2 B31110	conserved hypothet
26	147	15.2	375	2 B98177	hypothetical prote
27	145.5	15.1	654	2 B71623	knob-associated Hi
28	145.5	15.1	1398	2 T13741	hypothetical prote
29	144.5	15.0	60	2 C64698	probable histidine

30	143.5	14.9	852	2 A34373	histidine-rich cal
31	143	14.8	731	1 JC2464	probable copper-tr
32	143	14.8	940	1 JE0291	FB19 protein - hum
33	142.5	14.8	102	2 T30119	hypothetical prote
34	142.5	14.8	270	2 A26480	knob protein - mal
35	142.5	14.8	473	2 A54494	knob-associated hi
36	142.5	14.8	634	2 A28412	histidine-rich pro
37	142.5	14.8	634	2 A54495	knob protein precu
38	142	14.7	301	2 S31782	HRP11 protein - ma
39	142	14.7	997	2 T28872	hypothetical prote
40	141	14.6	905	2 AC2680	hypothetical prote
41	141	14.6	916	2 A97462	hypothetical prote
42	140	14.5	297	2 S21737	proline-rich prote
43	140	14.5	309	2 A25942	histidine/alanine-
44	139	14.4	130	2 S14983	extensin class I (
45	139	14.4	330	2 S22140	nodulin Enod2 - Se

ALIGNMENTS

RESULT 1

KGHUGH
histidine-rich glycoprotein precursor - human
N:Alternate names: HRG
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C:Accession: A01287; S29669
R:Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.
Biochemistry 25, 2220-2225, 1986
A:Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucleolus
A:Reference number: A01287; MUID:86216149; PMID:3011081
A:Accession: A01287
A:Molecule type: mRNA
A:Residues: 1-525 <KOI>
A:Cross-references: UNIPROT:P04196; GB:AB005803; NID:g2280513; PIDN:BAA21613.1; PID:g2280513
R:Hennis, B.; Havelaar, A.; Klufft, C.
Submitted to the EMBL Data Library, October 1991
A:Description: PCR detection of a dinucleotide repeat in the human histidine-rich glycoprotein
A:Reference number: S29669
A:Accession: S29669
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 214-247 <HEN>
A:Cross-references: EMBL:217218; NID:g32453; PIDN:CAA78925.1; PID:g32454
C:Comment: Although its physiological function is not yet known, HRG does bind heme, dyes and the lysine-binding site of plasminogen. On the basis of its homology with HMW knob protein, it is thought to be involved in blood coagulation cascade.
C:Comment: The amino half of this protein is homologous to the first two cystatin-like domain, could not have inhibitory activity.
C:Comment: In addition to having a high histidine and proline content, this protein has a high e-rich' region.
C:Genetics:
A:Gene: GDB:HRG
A:Cross-references: GDB:120055; OMIM:142640
A:Map position: 3q27-3q27
C:Superfamily: histidine-rich glycoprotein; cystatin homology
C:Keywords: duplication; glycoprotein; heparin binding; tandem repeat
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-525/Product: histidine-rich glycoprotein #status predicted <MAT>
F:19-131/Domain: cystatin homology <CY1>
F:140-246/Domain: cystatin homology <CY2>
F:276-321/Region: proline-rich
F:348-437/Region: histidine-rich
F:351-497/Region: proline-rich
F:63,125,344,345/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:78-89,105-126,218-241/Disulfide bonds: #status predicted

Query Match 100.0%; Score 966; DB 1; Length 525;
Best Local Similarity 100.0%; Pred. NO. 1.4e-74;
Matches 148; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HPHKHSHEQHGHGHHPHAHPHEDTHRQHPGHGHHPHGHGHHPHGHCH 60

Db 350 HPHKHSHEQHPRGHHPHAAHPHEHDTHQHPGHHPHGHHPHGHHPGHHPGHHPCH 409

QY 61 DFQDYGCDPPPHNQHCCHGCPPPGHLLRRRGKPPPHCRQIGSYVYRLPPLRKGEV 120

Db 410 DFQDYGCDPPPHNQHCCHGCPPPGHLLRRRGKPPPHCRQIGSYVYRLPPLRKGEV 469

QY 121 LPLPEANFSPFLPHHKHPLKPDNQPPF 148

Db 470 LPLPEANFSPFLPHHKHPLKPDNQPPF 497

RESULT 2

A60488

histidine-rich glycoprotein - bovine (fragments)

N:Alternate names: autorosette inhibition factor

C:Species: Bos primigenius taurus (cattle)

C>Date: 19-Mar-1993 #sequence revision 23-Mar-1995 #text_change 07-Jul-1995

C:Accession: S35687; JG2196; A60488

R:Sorensen, C.B.; Krogh-Pedersen, H.; Petersen, T.E.

FEBS Lett. 328, 285-290, 1993

A:Title: Determination of the disulphide bridge arrangement of bovine histidine-rich gly

A:Reference number: S35687; MUID:93351678; PMID:8348977

A:Accession: S35687

A:Molecule type: protein

A:Residues: 1-25;26-52;57-75;82-88;95-119;146-173;175-206;210-309;313-445 <SOR>

A>Note: 355-Gln and 368-Tyr were also found

R:Halkier, T.; Andersen, H.; Vestergaard, A.; Magnusson, S.

Biochem. Biophys. Res. Commun. 200, 78-82, 1994

A:Title: Bovine histidine-rich glycoprotein is a substrate for bovine plasma factor XIII

A:Reference number: JC2196; MUID:94220160; PMID:7909439

A:Accession: JC2196

A:Molecule type: protein

A:Residues: 1-23;35-54,'VK',57-101,'R','TVGEYTEG',116,'N',118,'R',120-136;137-145;150-20

A:Experimental source: plasma

R:Vestergaard, A.B.; Andersen, H.F.; Magnusson, S.; Halkier, T.

Thromb. Res. 60, 385-396, 1990

A:Title: Histidine-rich glycoprotein inhibits contact activation of blood coagulation.

A:Reference number: A60488; MUID:91196010; PMID:2084959

A:Accession: A60488

A:Molecule type: protein

A:Residues: 1-6,'X',8-15 <VES>

C:Comment: This protein is a single-chained plasma protein which participates in transgl

C:Superfamily: histidine-rich glycoprotein; cystatin homology

C:Keywords: glycoprotein; plasma

F:2-113/Domain: cystatin homology <CV1>

F:122-207/Domain: cystatin homology (fragments) <CY2>

F:7-424,60-71,87-108,165-346,180-203,258-288/Disulfide bonds: #status experimental

F:74,107,164,266/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 52.1%; Score 503.5; DB 2; Length 445;

Best Local Similarity 49.2%; Pred. No. 1.7e-35;

Matches 91; Conservative 15; Mismatches 26; Indels 53; Gaps 6;

QY 1 HPHKHS-----HEQH-----PH-----PH-----GHH--PHAHHPH 23

Db 249 HPESHVNRCPPELEKHNDSPPFOARAPLPPFPGLRCPPFPFGTKGNHRPPHDHSD 308

QY 24 ENDTHRQHPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 83

Db 309 EH-----HNHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 352

QY 84 PPGHLRRRGKPPPHCRQIGSYVYRLPPLRKGEVLPLPEANFSPFLPHHKHPLKPD 143

Db 353 PPRHSKRGCPKGFHFRFWRFTGTIHLPLSLKKGEVLPLPEANFSPFLPHNNPLQPE 412

QY 144 NQFPF 148

Db 413 IQAFP 417

histidine-rich glycoprotein precursor - Plasmodium lophurae

C:Species: Plasmodium lophurae

C>Date: 30-Sep-1987 #sequence_revision 30-Sep-1987 #text_change 09-Jul-2004

C:Accession: A22692

R:Kavetch, J.V.; Feder, R.; Pavlovic, A.; Blobel, G.

Nature 312, 616-620, 1984

A:Title: Primary structure and genomic organization of the histidine-rich protein of the

A:Reference number: A22692; MUID:85061618; PMID:6095114

A:Accession: A22692

A:Molecule type: DNA

A:Residues: 1-351 <RAV>

A:Cross-references: UNIPROT:P04929; GB:X01469; NID:99997; PIDN:CAA25698.1; PID:g99999

C:Comment: There are two copies of 16-residue repeats, two copies of 17-residue repeats,

C:Genetics:

A:Introns: 23/3

C:Superfamily: plasmodium histidine-rich protein

C:Keywords: glycoprotein; tandem repeat

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-47/Domain: propeptide #status predicted <PRO>

F:48-351/Product: histidine-rich glycoprotein #status predicted <MAT>

F:59-74,75-90/Region: 16-residue repeats

F:91-107,108-123/Region: 17-residue repeats

F:124-136,139-153/Region: 15-residue repeats

F:173-301,312-331/Region: 10-residue repeats

F:40/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 25.0%; Score 241.5; DB 1; Length 351;

Best Local Similarity 47.7%; Pred. No. 2e-13; 34; Indels 11; Gaps 3;

Matches 42; Conservative 1; Mismatches 1; Indels 11; Gaps 3;

QY 2 PHKHSHEQHHPGHHPH-----PHAHHPHEH--DTHRQHPGHHPHGHHPHGHHPHGH 54

Db 184 PHHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHH 243

QY 55 HPHCHDQDYGCDPPPHNQHCCHG 82

Db 244 HHHHHHHHD----AHHHHHHHHDAAHH 267

RESULT 4

A54523

histidine-rich protein - Plasmodium lophurae (fragment)

C:Species: Plasmodium lophurae

C>Date: 15-Oct-1994 #sequence_revision 15-Oct-1994 #text_change 09-Jul-2004

C:Accession: A54523

R:Irving, D.O.; Cross, G.A.M.; Feder, R.; Wallach, M.

Mol. Biochem. Parasitol. 18, 223-234, 1986

A:Title: Structure and organization of the histidine-rich protein gene of Plasmodium loph

A:Reference number: A54523; MUID:86174893; PMID:3007981

A:Accession: A54523

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-140 <IRV>

A:Cross-references: UNIPROT:Q26056; GB:M15317; NID:g160331; PIDN:AAA29616.1; PID:g552196

C:Superfamily: plasmodium histidine-rich protein

C:Keywords: tandem repeat

Query Match 22.5%; Score 217; DB 2; Length 140;

Best Local Similarity 39.6%; Pred. No. 1e-11;

Matches 38; Conservative 2; Mismatches 3; Indels 22; Gaps 2;

QY 2 PHKHSHEQHHPGHHPHHPHHPHHPHHPHHPHHPHHPHHPHHPHHPHHPHHPH 46

Db 8 PHHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHHHPHHHHHP 67

QY 47 PHGHHHPGHHPCHDFQDYGCDPPPHNQHCCHG 82

Db 68 HHHHPWFHHHHHDAAHH-----HHHHHDAAHH 96

RESULT 5

T45059

hypothetical protein Y39B6B.gg [imported] - Caenorhabditis elegans

C:Species: Caenorhabditis elegans
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C:Accession: T45059
R:Wilson, R.; Alincough, R.; Anderson, K.; Baynes, C.; Berks, M.; Bonfield, J.; Burton, R.; raser, A.; Fulton, L.; Gardner, A.; Green, P.; Hawkins, T.; Hillier, L.; Jier, M.; Johns, B.; O'Callaghan, M.; Parsons, J.; Percy, C.; Rifken, L.; Roopra, A.; Saunders, D. Nature 368, 32-38, 1994
A:Authors: Shownkeen, R.; Sims, M.; Smaldon, N.; Smith, A.; Smith, M.; Sonnhammer, E.; Stock, L.; Wilkinson-Sproat, J.; Wohlman, P.
A:Title: 2.2 Mb of contiguous nucleotide sequence from chromosome III of C. elegans.
A:Reference number: S43531; MUID:94150718; PMID:7906398
A:Accession: T45059
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-735 <WIL>
A:Cross-references: UNIPROT:Q9NES7; EMBL:AL132896; NID:G6434440; PIDN:CAB60938.1; PID:96
A:Experimental source: clone Y39B6B
C:Genetics:
A:Map position: 3
A:Introns: 18/1; 69/1
A:Note: Y39B6B.99

Query Match 20.0%; Score 193; DB 2; Length 735;
Best Local Similarity 27.9%; Pred. No. 5.2e-09;
Matches 60; Conservative 4; Mismatches 65; Indels 86; Gaps 11;

QY 1 HPKHHS-HEQHPGHHPHNAHPH-----PHGHH-----PHGHH-----PHCHDF 62
DB 434 HAPAHGHGHSCHGHSHPAHGHGHHNAHPAHGHGHHNAHPAHGHGHHGHTHHGH 493
QY 32 PHGHPH--HGHH-----PHGHH-----PHGHH-----PHCHDF 62
DB 494 HGSHPHHPHGHGHHNAHPAHGHGHHGHTHHGHGHHNAHPAHGHGHHGHTHHGHG 553
QY 63 QDYGPC-----DPPPHNQHCCHG-HGPPPGHLRRRGPG-----KGPRPFC 103
DB 554 SHSPAHGHGHHNAHPAHGHGHHGHSVHGHSHGHHNAHPAHGHGHHGVH 613

QY 104 ROIGSVYRLPPKRGVLPPEANFSPFLPHKH 138
DB 614 GHGAGYGAHHGHGA-----HHHNAHPHHEH 639

RESULT 6
T14155
zinc finger protein Peg3 - mouse
C:Species: Mus musculus (house mouse)
C:Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C:Accession: T14155
R:Kuroiwa, Y.; Kaneko-Ishino, T.; Kagitani, F.; Kohda, T.; Li, L.L.; Tada, M.; Suzuki, R. Nature Genet. 12, 186-190, 1996
A:Title: Peg3 imprinted gene on proximal chromosome 7 encodes for a zinc finger protein.
A:Reference number: Z17892; MUID:96154192; PMID:8563758
A:Accession: T14155
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1571 <KUR>
A:Cross-references: UNIPROT:O54978; EMBL:AF038939; NID:G2791677; PID:G2791678; PIDN:AAB9
C:Genetics:
A:Gene: Peg3
A:Map position: 7
C:Keywords: zinc finger

Query Match 18.2%; Score 176; DB 2; Length 1571;
Best Local Similarity 36.5%; Pred. No. 2.9e-07;
Matches 31; Conservative 5; Mismatches 45; Indels 4; Gaps 1;

QY 2 PHKHSHSQHPGHHPHNAHPHEDTHRQHPGHHPGHHPGHHPGHHPGHHPCHD 61
DB 977 PLVQEMRSEEPHDDKPHGQEPHDDKPHGQEPHDDKPHGQEPHDDKPHGQEPHDDK 1036
QY 62 FQDYGCDPPPHNQHCCHGHPPP 86

Db 1037 PIDQMRSEEPHSE-----ESHGDEP 1057

RESULT 7
T33997
hypothetical protein W03G1.5 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T33997
R:Pauley, A.; Scheet, P.; Harper, M.
A:Submitted to the EMBL Data Library, February 1999
A:Description: The sequence of C. elegans cosmid W03G1.
A:Reference number: Z21454
A:Accession: T33997
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-471 <PAU>
A:Cross-references: UNIPROT:Q9UAY0; EMBL:AF125964; PIDN:AAD14753.1; GSPDB:GN00022; CESP:
A:Experimental source: strain Bristol N2; clone W03G1
C:Genetics:
A:Gene: CESP:W03G1.5
A:Map position: 4

Query Match 18.0%; Score 173.5; DB 2; Length 471;
Best Local Similarity 40.6%; Pred. No. 1.5e-07;
Matches 39; Conservative 2; Mismatches 36; Indels 19; Gaps 6;

QY 3 HKHSHSQHPGH-----HPHAPHPEHEDTHRQHPGHHPGHHPGHHP-----HGHHHPGH 51
DB 369 HGSRSHPSPRGHSGRHPGPHCPGRHG-HGPPPHHHHDGRSPSRGHHHHHHHGCRRPP 427
QY 52 PHGHPHCHDFQDYGPCDP-----PPHQHCCHG 82
DB 428 PHHGHHHPFPF--WPFCPPPPFPWPPHRRGGCHHH 461

RESULT 8
A29454
knob-associated histidine-rich protein precursor - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C:Accession: A29454
R:Triglia, T.; Stahl, H.D.; Crewther, P.E.; Scanlon, D.; Brown, G.V.; Anders, R.F.; Kemp, EMBO J. 6, 1413-1419, 1987
A:Title: The complete sequence of the gene for the knob-associated histidine-rich protein
A:Reference number: A29454; MUID:87275836; PMID:3301325
A:Accession: A29454
A:Molecule type: DNA
A:Residues: 1-657 <TRI>
A:Cross-references: UNIPROT:P06719; GB:Y00060; NID:G9908; PIDN:CAA68268.1; PID:G9909
C:Superfamily: knob-associated histidine-rich protein

Query Match 17.8%; Score 171.5; DB 2; Length 657;
Best Local Similarity 39.0%; Pred. No. 3.1e-07;
Matches 32; Conservative 5; Mismatches 32; Indels 13; Gaps 2;

QY 7 SHEQHPGHHPHAPHPEHEDTHRQHPGHHPGHHPGHHP-----HG-----HHHP 53
DB 57 AQKQHEHHHHHHHHHHHHQHAPHQAPHAHHHHHGEVNHQAPVHQVHQDQAHHHH 116
QY 54 GHHPHCHDFQDYGPCDPPPHNQ 75
DB 117 HHHHHLHPQPGQGTVANPPSNE 138

RESULT 9
H83994
hypothetical protein BH2760 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: H83994
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Maui, N.; Fuji, F.; Hiran Nucleic Acids Res. 28, 4317-4331, 2000

RESULT 13

T23089

hypothetical protein H13N06.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

C:Accession: T23089

R:Jennard, N.

submitted to the EMBL Data Library, October 1997

A:Reference number: Z19673

A:Accession: T23089

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-515 <MIL>

A:Cross-references: UNIPROT:Q9XT07; EMBL:Z99942; PIDN:CAB17070.1; GSPDB:GN00028; CESP:HL

A:Experimental source: clone H13N06

C:Genetics:

A:Gene: CESP:H13N06.5

A:Map position: X

A:introns: 118/1; 156/2; 182/2; 306/1

Query Match

Best Local Similarity 17.2%; Score 166.5; DB 2; Length 515;

Matches 39; Conservative 6; Mismatches 29; Indels 63; Gaps 7;

QY

5 HHSHEQHPHGHPHAAH-----HPHEHDTTHRQHPHGHPHG 39

DB 92 HHQGHGHAHGHHADADGGCPYAKAAAAAATAAHDGHAHDHGHADHDG-HAHD 150

QY

40 HHPHGH-----HPHGHPHGHPCHDFQDYGCPDPPPHNQHCCHGCHLRRGP 94

DB 151 HHGHSDEEDHHHG-HAHDHGHSH--EDHG-----HSHGAE----- 185

QY

95 GKGRPFHCRQIGSVYR 111

DB 186 -----SAKQVGDEYQ 195

RESULT 14

T30173

zinc finger protein Pw1 - mouse

C:Species: Mus musculus (house mouse)

C>Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 09-Jul-2004

C:Accession: T30173

R:Relaix, F.; Weng, X.; Marazzi, G.; Yang, E.; Copeland, N.; Jenkins, N.; Spence, S.E.;

Dev. Biol. 177, 383-396, 1996

A:Title: Pw1, a novel zinc finger gene implicated in the myogenic and neuronal lineages.

A:Reference number: Z20754; MUID:96400442; PMID:8806818

A:Accession: T30173

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: DNA

A:Residues: 1-1378 <REL>

A:Cross-references: UNIPROT:O61138; EMBL:U48804; NID:g1197856; PID:g1197857; PIDN:AAC527

A:Experimental source: strain Sv129; limb bud

C:Genetics:

A:Map position: 7

Query Match

Best Local Similarity 17.2%; Score 166.5; DB 2; Length 1378;

Matches 31; Conservative 5; Mismatches 45; Indels 5; Gaps 2;

QY

2 PHKHSHSEQHPGHHPHAEH-DTHRQHPGHHPGHHPGHHPGHHPGHHPGH 60

DB 783 PLDQEMRSEPHDDKPHGQEPHDDMRPHGQEPHDDPHGQEPHGDPEPHDX 842

QY

61 DFQDYGCPDPPPHNQHCCHGHP 86

DB 843 EPIDQEMRSEPHSE-----ESHGDEP 864

RESULT 15

I49714

MHC H-2K/t-w5-linked ORF precursor - mouse

C:Species: Mus musculus (house mouse)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999

C:Accession: I49714

R:Han, S.

Mol. Cell. Biol. 10, 138-145, 1990

A:Title: A putative transmembrane protein with histidine-rich charge clusters encoded in

A:Reference number: I49714; MUID:90097821; PMID:2294398

A:Accession: I49714

A:Status: preliminary; translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-436 <RES>

A:Cross-references: GB:M32010; NID:g193738; PIDN:AAA37767.1; PID:g309286

Query Match

Best Local Similarity 16.8%; Score 162; DB 2; Length 436;

Matches 39; Conservative 2; Mismatches 32; Indels 18; Gaps 8;

QY

1 HPHKH-HSHEQHPGH-HPHAHHPHEHDTTHRQHPGH-----HPKH-HPHGHHPGHHPGH 53

DB 43 HGHSHGHSHEDFHGHSHGHSH-----EDFHGHGHGHSHGHSHGHSHGHSHGHSHGHSH 97

QY

54 GH-HPHCHDFQDYGCPDPPPHNQHCCHG 83

DB 98 GHSHGHSHDSLHHG-----GHGHAREHSHG 123

Search completed: June 15, 2005, 13:51:55

Job time : 30.5301 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:32:29 ; Search time 133.141 Seconds
(without alignments)
569.231 Million cell updates/sec

Title: US-10-074-225A-5

Perfect score: 966

Sequence: 1 HPHKHSHEQHGHGHPHAH.....PSFPLPHKHGLKPDNQPP 148

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	966	100.0	525	1	HRG_HUMAN	P04196 homo sapien
2	949	98.2	312	2	Q68DR3	Q68dr3 homo sapien
3	645.5	66.8	525	2	Q99PS5	Q99ps5 mus musculus
4	645.5	66.8	525	2	Q99PS6	Q99ps6 mus musculus
5	645.5	66.8	525	2	Q99PSB3	Q99psb3 mus musculus
6	645.5	66.8	536	2	Q6YK32	Q6yK32 mus musculus
7	645.5	66.8	536	2	Q6YKA2	Q6yKa2 mus musculus
8	644	66.7	525	2	Q99PS8	Q99ps8 rattus norv
9	644	66.7	546	2	Q9BGU1	Q9bgul bos taurus
10	636	64.8	515	2	Q99PS7	Q99ps7 rattus norv
11	614.5	63.6	510	2	Q9ESB2	Q9esb2 rattus norv
12	543	56.2	526	1	HRG_BOVIN	Q28640 oryctolagus
13	489	50.6	396	2	Q8MPF30	P33433 bos taurus
14	281.5	29.1	233	2	Q8MPF30	Q8mpf30 dictyosteli
15	241.5	25.0	351	1	HRXP_PLALO	P04929 plasmodium
16	223.5	23.1	294	2	Q9KI87	Q9ki87 bacillus ce
17	217	22.5	140	2	Q26056	Q26056 plasmodium
18	216.5	22.4	296	2	Q9KI89	Q9ki89 bacillus ce
19	213.5	22.1	465	2	Q801E5	Q801e5 xenopus lae
20	208	21.5	293	2	Q632V3	Q632v3 bacillus ce
21	206	21.3	248	2	Q6HCE3	Q6hce3 bacillus th
22	205.5	21.3	273	2	Q9KI88	Q9ki88 bacillus ce
23	205.5	21.3	308	2	Q72Y20	Q72y20 bacillus ce
24	205.5	21.3	462	2	Q7SVH2	Q7svh2 xenopus lae
25	205.5	21.3	462	2	Q7ZY91	Q7zy91 xenopus lae
26	205	21.2	295	2	Q816T1	Q816t1 bacillus ce
27	204.5	21.2	479	2	Q8T0B3	Q8t0b3 drosophila
28	204.5	21.2	783	2	Q9V5Z7	Q9v5z7 drosophila
29	202.5	21.0	430	2	Q6NBF4	Q6nbf4 rhodopseudo
30	199.5	20.7	137	2	Q9KI77	Q9ki77 bacillus an
31	199.5	20.7	133	2	Q9KI76	Q9ki76 bacillus an

32	199.5	20.7	265	2	Q9KI91	Q9ki91 bacillus an
33	198.5	20.5	474	2	Q6DIT5	Q6dit5 xenopus tto
34	196.5	20.3	449	1	CSUP_DROME	Q9v3a4 drosophila
35	193	20.0	735	2	Q9NES7	Q9nes7 caenorhabdi
36	191.5	19.8	341	2	Q9VV44	Q9vv44 drosophila
37	191.5	19.8	356	2	Q7KIP4	Q7kip4 drosophila
38	191.5	19.8	485	2	Q76876	Q76876 drosophila
39	190.5	19.7	1131	2	Q7SDJ5	Q7sdj5 ashbya gos
40	188	19.5	403	2	Q98H88	Q98h88 rhizobium l
41	183	18.9	213	2	Q6E104	Q6e104 cucurbita m
42	182	18.8	118	2	Q9KI82	Q9ki82 bacillus an
43	182	18.8	121	2	Q9KI80	Q9ki80 bacillus an
44	182	18.8	124	2	Q9KI78	Q9ki78 bacillus an
45	182	18.8	124	2	Q9KI79	Q9ki79 bacillus an

ALIGNMENTS

RESULT 1

HRG_HUMAN	STANDARD;	PRT;	525 AA.
ID			
AC	P04196;		
DT	20-MAR-1987 (Rel. 04, Created)		
DT	20-MAR-1987 (Rel. 04, Last sequence update)		
DT	05-JUL-2004 (Rel. 44, Last annotation update)		
DE	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG).		
DE	Name=HRG;		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
OX	NCBI_TaxID=9606;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	TISSUE=Liver;		
RX	MEDLINE=86216149; PubMed=3011081;		
RA	Koide T., Foster D.C., Yoshitake S., Davie E.W.;		
RT	"Amino acid sequence of human histidine-rich glycoprotein derived from the nucleotide sequence of its cDNA.";		
RL	Biochemistry 25:2220-2225(1986).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RA	Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;		
RL	Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE OF 214-247 FROM N.A.		
RX	MEDLINE=94245171; PubMed=8188234;		
RA	Hennis B.C., Franta R.R., Bakker E., Voseen R.H., van der Poort E.W.,		
RA	Blonden L.A., Cox S., Khan P.M., Spurr N.K., Kluff C.;		
RT	"Evidence for the absence of intron H of the histidine-rich glycoprotein (HRG) gene: genetic mapping and in situ localization of HRG to chromosome 3q28-q29.";		
RL	Genomics 19:195-197(1994).		
RN	[4]		
RP	SEQUENCE OF 19-27.		
RC	TISSUE=Plasma;		
RX	MEDLINE=93092337; PubMed=1459097;		
RA	Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,		
RA	Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,		
RA	Hochstrasser D.F.;		
RT	"Plasma protein map: an update by microsequencing.";		
RL	Electrophoresis 13:707-714(1992).		
CC	-!- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kininogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.		
CC	-!- SUBCELLULAR LOCATION: Secreted.		
CC	-!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.		
CC	-!- DOMAIN: In addition to having a high His and Pro content, this		


```
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055898; BAB33095.1; -.
DR EMBL; BC011168; AAH11168.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59162 MW; A83B93A439CFB126 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHSHEQHPGHHPHAEHDTTHQHPGHGHHHPGHGHHHPGHGHHHPGHGHHHPCH 60
D 347 HPHSHHPGHHSHGHHHPGHSHSHGHHHPGHGHHHPGHGHHHPGHGHHHPGHGHH 406
QY 61 DFQDYGCDPPPHNQ---GHCCHGHPGPHGHLRRRGKGPFPFHCRCQIGSVYRLPLRK 117
D 407 DFLDYGCDPPSPNSQELKGQYHRGYPGPHGSHRKGKGLFPFHQQIGYVYRLPLNI 466
QY 118 GEVLPLPEANFPSPFLPHKHPLKPNQPPF 148
D 467 GEVLTLPEANFPSPFLPNCNRSLOPEIQPFP 497

RESULT 4
ID Q99PS6 PRELIMINARY; PRT; 525 AA.
AC Q99PS6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg; Synonyms=MMHRG;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055897; BAB33094.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59090 MW; A83B93A439CFB3AC CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHSHEQHPGHHPHAEHDTTHQHPGHGHHHPGHGHHHPGHGHHHPGHGHHHPCH 60
D 347 HPHSHHPGHHSHGHHHPGHSHSHGHHHPGHGHHHPGHGHHHPGHGHHHPGHGHH 406
QY 61 DFQDYGCDPPPHNQ---GHCCHGHPGPHGHLRRRGKGPFPFHCRCQIGSVYRLPLRK 117
D 407 DFLDYGCDPPSPNSQELKGQYHRGYPGPHGSHRKGKGLFPFHQQIGYVYRLPLNI 466
QY 118 GEVLPLPEANFPSPFLPHKHPLKPNQPPF 148
D 467 GEVLTLPEANFPSPFLPNCNRSLOPEIQPFP 497

RESULT 5
ID Q99SB3 PRELIMINARY; PRT; 525 AA.
AC Q99SB3;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Parish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
cellular origin.";
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194028; AAG28416.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59132 MW; 6E55F2A439CFB123 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 525;
Best Local Similarity 69.5%; Pred. No. 1.2e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHSHEQHPGHHPHAEHDTTHQHPGHGHHHPGHGHHHPGHGHHHPGHGHHHPCH 60
D 347 HPHSHHPGHHSHGHHHPGHSHSHGHHHPGHGHHHPGHGHHHPGHGHHHPGHGHH 406
QY 61 DFQDYGCDPPPHNQ---GHCCHGHPGPHGHLRRRGKGPFPFHCRCQIGSVYRLPLRK 117
D 407 DFLDYGCDPPSPNSQELKGQYHRGYPGPHGSHRKGKGLFPFHQQIGYVYRLPLNI 466
QY 118 GEVLPLPEANFPSPFLPHKHPLKPNQPPF 148
D 467 GEVLTLPEANFPSPFLPNCNRSLOPEIQPFP 497

RESULT 6
ID Q6YK32 PRELIMINARY; PRT; 536 AA.
AC Q6YK32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Histidine-rich glycoprotein Hrg.
GN Name=Hrg;
OS Mus musculus (Mouse);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137504; AAN27996.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60439 MW; CCF14FC08DD53D7C CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 536;
Best Local Similarity 69.5%; Pred. No. 1.3e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;
```

```
QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 60
DQ 358 HPHSHHPGHSHGHHPHGHSHSHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 417
QY 61 DFQDYGCPDPPPHNQ---GHCCHGCPDPPHARRGPGKPRPFHRCQIGSVYRLPLRK 117
DQ 418 DFLDYGCPDPPSNSQBLKQYHRGYPGPHGHGRKPGKGLPFFHQQIGYVYRLPLNI 477
QY 118 GEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508

RESULT 7
Q6YKA2
ID Q6YKA2 PRELIMINARY; PRT; 536 AA.
AC Q6YKA2;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Histidine-rich glycoprotein Hrg.
GN Name=Hrg1;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NIH/Ola.
RA Hsu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY135662; AA10183.1; -.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot_inh_cysstat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
DR SEQUENCE 536 AA; 6492 MW; 71CE4FC6DF3A3D72 CRC64;

Query Match 66.8%; Score 645.5; DB 2; Length 536;
Best Local Similarity 69.5%; Pred. No. 1.3e-42;
Matches 105; Conservative 9; Mismatches 34; Indels 3; Gaps 1;

QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 60
DQ 358 HPHSHHPGHSHGHHPHGHSHSHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 417
QY 61 DFQDYGCPDPPPHNQ---GHCCHGCPDPPHARRGPGKPRPFHRCQIGSVYRLPLRK 117
DQ 418 DFLDYGCPDPPSNSQBLKQYHRGYPGPHGHGRKPGKGLPFFHQQIGYVYRLPLNI 477
QY 118 GEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508

RESULT 8
Q99PS8
ID Q99PS8 PRELIMINARY; PRT; 525 AA.
AC Q99PS8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein 1.
GN Name=RNHRG1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;

QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 60
DQ 358 HPHSHHPGHSHGHHPHGHSHSHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 417
QY 61 DFQDYGCPDPPPHNQ---GHCCHGCPDPPHARRGPGKPRPFHRCQIGSVYRLPLRK 117
DQ 418 DFLDYGCPDPPSNSQBLKQYHRGYPGPHGHGRKPGKGLPFFHQQIGYVYRLPLNI 477
QY 118 GEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508
DQ 478 GEVLPLPEANFFSPLPNCNRSLOPEIQPFP 508

RESULT 9
Q9BGUL
ID Q9BGUL PRELIMINARY; PRT; 546 AA.
AC Q9BGUL;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=BTHRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055894; BAB33091.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cysstat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
DR SEQUENCE 546 AA; 61948 MW; 26264858824D89EE CRC64;

Query Match 66.7%; Score 644; DB 2; Length 546;
Best Local Similarity 62.7%; Pred. No. 1.7e-42;
Matches 104; Conservative 10; Mismatches 34; Indels 18; Gaps 2;

QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 60
DQ 353 HPHGHHPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 412
QY 61 -----DFQDYGCPDPPHN---QHCCHGCPDPPHARRGPGKPRPFH 102
DQ 413 HPGHHPHGHHPHNDNDYDHGCDPPHRRQDPPHRRQDPPHRRQDPPHRRQDPPHRR 472
QY 103 CRQIGSVYRLPLRKGEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 473 WRPTGYIHLPSLKKGEVLPLPEANFFSPLPNCNRSLOPEIQPFP 518
DQ 473 WRPTGYIHLPSLKKGEVLPLPEANFFSPLPNCNRSLOPEIQPFP 518

RESULT 10
Q99PS7
ID Q99PS7 PRELIMINARY; PRT; 515 AA.
AC Q99PS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
```

```
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055895; BAB33092.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cysstat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
DR SEQUENCE 525 AA; 59049 MW; 38290A631FAC7777 CRC64;

Query Match 66.7%; Score 644; DB 2; Length 525;
Best Local Similarity 68.6%; Pred. No. 1.6e-42;
Matches 107; Conservative 7; Mismatches 34; Indels 8; Gaps 2;

QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 55
DQ 342 HPHGHHPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 401
QY 56 HPHCHDFQDYGCPDPPHNQ---GHCCHGCPDPPHARRGPGKPRPFHRCQIGSVYRL 112
DQ 402 HPHGHDFLDYGCDDPPSNSQBLKQYHRGYPGPHGHGRKPGKGLPFFHQQIGYVYRL 461
QY 113 PPLRKGEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 462 PPLNVEVLTPPEANFFSPLPNCNRSLOPEIQPFP 497
DQ 462 PPLNVEVLTPPEANFFSPLPNCNRSLOPEIQPFP 497

RESULT 9
Q9BGUL
ID Q9BGUL PRELIMINARY; PRT; 546 AA.
AC Q9BGUL;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=BTHRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055894; BAB33091.1; -.
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cysstat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
DR SEQUENCE 546 AA; 61948 MW; 26264858824D89EE CRC64;

Query Match 66.7%; Score 644; DB 2; Length 546;
Best Local Similarity 62.7%; Pred. No. 1.7e-42;
Matches 104; Conservative 10; Mismatches 34; Indels 18; Gaps 2;

QY 1 HPHKHSHHQPHGHHPHHAHPHEHDTHQPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 60
DQ 353 HPHGHHPGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 412
QY 61 -----DFQDYGCPDPPHN---QHCCHGCPDPPHARRGPGKPRPFH 102
DQ 413 HPGHHPHGHHPHNDNDYDHGCDPPHRRQDPPHRRQDPPHRRQDPPHRRQDPPHRR 472
QY 103 CRQIGSVYRLPLRKGEVLPLPEANFFSPLPHKHPKLPDNPQPP 148
DQ 473 WRPTGYIHLPSLKKGEVLPLPEANFFSPLPNCNRSLOPEIQPFP 518
DQ 473 WRPTGYIHLPSLKKGEVLPLPEANFFSPLPNCNRSLOPEIQPFP 518

RESULT 10
Q99PS7
ID Q99PS7 PRELIMINARY; PRT; 515 AA.
AC Q99PS7;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
```



```
FT CARBOHYD 115 115 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 192 192 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 240 240 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 310 310 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 485 485 N-linked (GlcNAc. . .) (Potential).
FT SITE 303 304 Cleavage (by plasmin).
FT SITE 421 422 Cleavage (by plasmin).
SQ SEQUENCE 526 AA; 58877 MW; 810F23D367D93D42 CRC64;

Query Match 56.2%; Score 543; DB 1; Length 526;
Best Local Similarity 57.6%; Pred. No. 1.2e-34;
Matches 98; Conservative 8; Mismatches 40; Indels 24; Gaps 4;

QY 2 PHKHSHSQPHGHHPHA-----HH-----PHEHDTHQPHGHHPGHHP 42
Db 330 PHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPHGHHPGPPPH 389
QY 43 HGHHPHGHHPHGHHPHCHFDQYGCDDPPHNOG-----HCCGHGPPPHGLRRRFGKGP 98
Db 390 HGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHP 448
QY 99 RPHCRQIGSVYRLPLRKGEVLPPEANFPSPFLPHHKKHPLKPDNOPEP 148
Db 449 FPFHMRIRGVSQVLPQKGEVLPPEANFPOLLRNHHTHPLKPEIQPPP 498

RESULT 13
ID_HRG_BOVIN STANDARD; PRT; 396 AA.
AC F33433;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein (Histidine-proline rich glycoprotein)
DE (HPRG) (Fragments).
GN Name=HRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE, AND DISULFIDE BONDS.
RX MEDLINE=93351678; PubMed=8348977; DOI=10.1016/0014-5793(93)80945-Q;
RA Soerensen C.B., Krogh-Pedersen H., Petersen T.E.;
RT "Determination of the disulphide bridge arrangement of bovine
RT histidine-rich glycoprotein."
RL FEBS Lett. 328:285-290(1993).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-
CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cysstatin-like domains.
DR InterPro: IPR000010; Prot_inh_cysstat.
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat.
FT DOMAIN 1 102 Cystatin-like 1.
FT DOMAIN 103 169 Cystatin-like 2.
FT DOMAIN 191 238 Pro-rich.
FT DOMAIN 243 368 His/Pro-rich.
FT DISULFID 7 375
FT DISULFID 56 67
FT DISULFID 77 92
FT DISULFID 123 297
FT DISULFID 137 160
FT DISULFID 212 242
```

```
FT NON CONS 52 53
FT CARBOHYD 70 70 N-linked (GlcNAc. . .).
FT NON CONS 71 72
FT NON CONS 78 79
FT CARBOHYD 91 91 N-linked (GlcNAc. . .).
FT NON CONS 103 104
FT CARBOHYD 122 122 N-linked (GlcNAc. . .).
FT NON CONS 163 164
FT CARBOHYD 220 220 N-linked (GlcNAc. . .).
FT NON CONS 263 264
FT NON CONS 303 304
FT VARIANT 86 86 S -> R.
FT VARIANT 309 309 S -> Q.
FT VARIANT 322 322 H -> Y.
SQ SEQUENCE 396 AA; 44470 MW; 128A8223499DB6FC CRC64;

Query Match 50.6%; Score 489; DB 1; Length 396;
Best Local Similarity 47.8%; Pred. No. 1.5e-30;
Matches 88; Conservative 16; Mismatches 26; Indels 54; Gaps 6;

QY 1 HPHKHHS-----HEQH-----PH-----GHHPHAHHPHE 24
Db 203 HPHSYNFRCPPLLEKHNHSDSPFOARAPLPFPFPPGLRCPHPFGTKGNHRP-----PHD 258
QY 25 HDTHRQHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHP 84
Db 259 HSSDE-----HHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHP 304
QY 85 PPHGLRRGPGKPPPHCRQIGSVYRLPLRKGEVLPPEANFPSPFLPHHKKHPLKPDN 144
Db 305 PPHSKERGPGKHGFRFHWRTGYIHRPLSLKKGEVLPPEANFPSPFLPHHKKHPLKPDN 144
QY 145 QPFP 148
Db 365 QAFP 368

RESULT 14
Q8MP30 PRELIMINARY; PRT; 233 AA.
AC Q8MP30; Q8T164;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein (similar to Plasmodium lophurae. Histidine-rich
DE glycoprotein).
OS Dictyostelium discoideum (Slime mold). Dictyostelium.
OC Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910; DOI=10.1038/nature00847;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumgart C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tungal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RA "Sequence and analysis of chromosome 2 of Dictyostelium discoideum."
RT Nature 418:79-85(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RA Baumgart C.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC123513; AAM44363.1; -.
DR EMBL; AC117070; AAM09303.2; -.
KW Hypothetical protein.
```


This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:31:29 ; Search time 95.7269 Seconds
(without alignments)
408.065 Million cell updates/sec

Title: US-10-074-225a-6

Perfect score: 697

Sequence: 1 SYNIIHRPPHGHGPPPP.....HPPHGHGFHDHGCDPPPSHK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Genesecp1980s:*
- 2: Genesecp1990s:*
- 3: Genesecp2000s:*
- 4: Genesecp2001s:*
- 5: Genesecp2002s:*
- 6: Genesecp2003as:*
- 7: Genesecp2003bs:*
- 8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	697	100.0	101	ADH10412	Rabbit HP
2	697	100.0	526	ABB79805	Rabbit hi
3	697	100.0	526	ADH10410	Rabbit HP
4	314	45.1	525	ABB79804	Human his
5	314	45.1	525	ADH76897	Human pro
6	314	45.1	525	ADH10409	Human HPR
7	314	45.1	525	ADH12335	Human ste
8	280.5	40.2	148	ABB79807	Rabbit hi
9	280.5	40.2	148	ABB79806	Human his
10	280.5	40.2	148	ADH10411	Human HPR
11	264.5	37.9	491	ABG09949	Novel hum
12	259.5	37.2	296	ADH45442	Rat Prote
13	259.5	37.2	296	ADH57105	Rat Prote
14	246	35.3	274	ADH47240	Rat Prote
15	246	35.3	274	ADH48703	Rat Prote
16	233.5	33.5	274	AAR04119	ORF2 of E
17	233.5	33.5	309	AAR04116	ORF1 of E
18	233.5	33.5	331	ADH45740	Human Pro
19	233.5	33.5	331	ABO52997	Human spl
20	233.5	33.5	424	ABM60186	Tumour-as
21	231.5	33.2	192	AAO30246	Human Sap
22	231.5	33.2	208	AAO30188	Human cp3
23	231	33.1	173	ADJ68263	Human hea
24	231	33.1	267	AAO30408	Human sec
25	229.5	32.9	124	AAW50192	Amino aci

ALIGNMENTS

RESULT 1

ID	ADH10412	standard; protein; 101 AA.
XX	ADH10412;	
XX	ADH10412;	
DT	11-MAR-2004	(first entry)
XX		
DE	Rabbit HPRG protein H/P rich domain.	
XX		
KW	Tpm; tropomyosin; antiangiogenic receptor;	
KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;	
KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;	
KW	antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;	
KW	angiogenesis; gene therapy; rabbit.	
OS	Oryctolagus cuniculus.	
XX		
PN	WO2003077872-A2.	
XX		
PD	25-SEP-2003.	
XX		
PF	17-MAR-2003; 2003WO-US008060.	
XX		
PR	15-MAR-2002; 2002US-0364047P.	
XX		
PA	(ATTE-) ATTENDON LLC.	
XX		
PI	McCrack K, Donate F, Juarez J, Mazar AP;	
XX		
DR	WPI; 2004-090604/09.	
XX		
PT	New tropomyosin-related antiangiogenic receptor polypeptide, useful for	
PT	inhibiting endothelial cell migration, invasion, proliferation or	
PT	angiogenesis, inducing endothelial cell apoptosis, or treating tumors or	
PT	cancer.	
XX		
PS	Claim 6; SEQ ID NO 24; 117pp; English.	
XX		
CC	The invention relates to an isolated tropomyosin (Tpm)-related	
CC	antiangiogenic receptor polypeptide or peptide, which is a fragment	
CC	of a full-length native Tpm protein expressed on the surface of	
CC	endothelial cells, or a variant of the fragment. It has a molecular mass	
CC	of about 17 kDa and corresponds in its sequence to, or is a variant of,	
CC	an internal fragment of a native Tpm isoform which is a binding site for	
CC	antiangiogenic polypeptide agents; The isolated antiangiogenic receptor	
CC	polypeptide, peptide or variant has substantially the same biochemical	
CC	activity of binding to the antiangiogenic polypeptide agents, as does the	

Adg15008	Human SEC
Aab71667	Human col
Aab71654	Human col
Abb68218	Drosophil
Abu11891	Human ABC
Abb61183	Drosophil
Abb61362	Drosophil
Abp73921	Candida a
Adg15007	Human SEC
Abb57176	Mouse isc
Abb43006	Human sec
Abb78538	Ser-Pro-P
Adn30407	Human sec
Adn43007	Human sec
Adn20587	Bacterial
Adn20587	Bacterial
Adn22380	Antigen m
Aar239650	A. thalia
Adm80808	Human CAD
Aab26153	Neospora

26	229	32.9	230	7	ADG15008
27	228.5	32.8	93	4	AAB71667
28	228.5	32.8	93	4	AAB71654
29	228.5	32.8	102	4	ABB68218
30	228	32.7	247	6	ABU11891
31	225	32.3	80	4	ABB61183
32	223	32.0	594	4	ABB61362
33	223	32.0	1130	5	ABP73921
34	222	31.9	272	7	ADG15007
35	221	31.7	1378	5	ABB57176
36	220	31.6	220	8	ADN43006
37	215.5	30.9	104	5	ABB78538
38	215	30.8	162	7	AAO30407
39	214.5	30.8	205	8	ADN43007
40	213.5	30.6	449	8	ADS23375
41	211	30.3	518	8	ADN20587
42	210	30.1	487	2	AAR22380
43	209.5	30.1	205	5	ABR39650
44	208.5	29.9	163	8	ADM80808
45	208	29.8	865	3	AAB26153

CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding, two chain
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a rabbit HPRG
 CC protein His-Pro (H/P) rich domain.
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 697; DB 8; Length 101;
 Best Local Similarity 100.0%; Pred. No. 1.9e-51;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 60
 DB 1 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 60
 QY 61 GHPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 101
 DB 61 GHPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 101
 RESULT 2
 ABB79805
 ID ABB79805 standard; protein; 526 AA.
 XX
 AC ABB79805;
 XX
 XX 25-NOV-2002 (first entry)
 DT
 XX Rabbit histidine proline rich glycoprotein.
 DE
 XX Histidine proline rich glycoprotein; HPRG; rabbit; antiangiogenic;
 KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
 KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumor;
 KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
 XX
 OS Oryctolagus cuniculus.
 XX
 FH Key Location/Qualifiers
 FT Domain 251..296
 FT /note="Proline-rich domain"
 FT Domain 321..421
 FT /note="Histidine-proline-rich domain, region also
 FT specifically claimed in Claim 1"
 XX
 PN WO200264621-A2.
 XX
 PD 22-AUG-2002.
 XX
 PF 14-FEB-2002; 2002WO-US004336.
 XX
 PR 14-FEB-2001; 2001US-0268370P.
 XX
 PA (ATTE-) ATTENUON LLC.
 XX
 PI Donate F, Harris S, Plunkett ML, Mazar AP;
 XX
 DR WPI; 2002-666989/71.
 DR P-PSDB; ABN84911.
 XX
 PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for

PT diagnosing or treating diseases associated with undesired cell migration,
 XX invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
 PS Claim 1; Page 12-13; 82pp; English.
 XX
 CC The present sequence is the protein sequence of rabbit histidine proline
 CC rich glycoprotein (HPRG), a proteinaceous chelator that can be used to
 CC inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
 CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
 CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
 CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
 CC endothelial cell proliferation or endothelial tube formation in vitro or
 CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
 CC Also claimed are: chemically synthesised or recombinantly produced
 CC peptide multimers; a diagnostically or therapeutically labeled anti-
 CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
 CC useful HPRG-related composition, comprising the diagnostically labeled
 CC polypeptide, peptide or peptide multimer and a carrier; an antibody
 CC specific for an epitope of HPRG that is present in the H/P domain of
 CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
 CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,
 CC or an antigen-binding fragment of the antibody; a method for inhibiting
 CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
 CC inducing apoptosis; a method for treating a subject having a disease or
 CC condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis; a method of stimulating or inhibiting
 CC angiogenesis in a subject; a method of detecting the presence of HPRG or
 CC its cleavage product or its peptide in a biological sample; isolated
 CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
 CC expression vector; transformed or transfected cells; a method of
 CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
 CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
 CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
 CC molecule or cells expressing the binding molecule, comprising the
 CC polypeptide, peptide or peptide multimer, immobilised to a solid support
 CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
 CC or enriching cells expressing HPRG-binding site or receptor, from a
 CC complex mixture. The compositions and methods are useful in diagnosing or
 CC treating a disease or condition associated with undesired cell migration,
 CC invasion, proliferation, or angiogenesis, such as cancer.
 CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
 CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
 CC inhibiting the growth of primary tumours or metastases, and may also be
 CC used in treating neurodegenerative diseases like Alzheimer's disease,
 CC Parkinson's disease and amyotrophic lateral sclerosis. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays
 XX
 SQ Sequence 526 AA;
 Query Match 100.0%; Score 697; DB 5; Length 526;
 Best Local Similarity 100.0%; Pred. No. 7.5e-51;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 60
 DB 321 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 380
 QY 61 GHPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 101
 DB 381 GHPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 421
 RESULT 3
 ADH10410
 ID ADH10410 standard; protein; 526 AA.
 XX
 AC ADH10410;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Rabbit HPRG protein sequence.

XX Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipeoriatic; dermatological; cardiant; vasotropic; vulnerary;
 KW angiogenesis; gene therapy; rabbit.
 XX
 OS Oryctolagus cuniculus.
 XX
 XX
 FH Key Location/Qualifiers
 FT Domain 251..296
 FT /note = pro-rich domain
 FT Domain 321..421
 FT /note = His-Pro rich domain
 FT
 XX WO2003077872-A2.
 PN
 XX
 XX
 PD 25-SEP-2003.
 XX
 XX 17-MAR-2003; 2003WO-US008060.
 XX
 PR 15-MAR-2002; 2002US-0364047P.
 XX
 XX (ATTE-) ATTENUON LLC.
 PA
 XX McCrae K, Donate F, Juarez J, Mazar AP;
 PI WPI; 2004-090604/09.
 XX
 DR New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 PT
 XX
 PS Claim 6; SEQ ID NO 22; 117pp; English.
 XX
 XX The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
 CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a rabbit HPRG
 CC protein sequence.
 XX
 SQ Sequence 526 AA;
 Query Match 100.0%; Score 697; DB 8; Length 526;
 Best Local Similarity 100.0%; Pred. No. 7.5e-51;
 Matches 101; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 60
 DB 321 SVNIHRPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 380

QY 61 GHPPHCPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 101
 DB 381 GHPPHCPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGH 421
 RESULT 4
 ABB79804
 ID ABB79804 standard; protein; 525 AA.
 XX
 AC ABB79804;
 XX
 DT 25-NOV-2002 (first entry)
 XX
 DE Human histidine proline rich glycoprotein.
 XX
 KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
 KW cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
 KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
 KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Domain 276..321
 FT /note = "proline-rich domain"
 FT Domain 350..497
 FT /note = "Histidine-proline-rich domain, region also
 FT specifically claimed in Claim 1"
 FT
 XX WO200264621-A2.
 PN
 XX
 PD 22-AUG-2002.
 XX
 PF 14-FEB-2002; 2002WO-US004336.
 XX
 PR 14-FEB-2001; 2001US-0268370P.
 XX
 PA (ATTE-) ATTENUON LLC.
 XX
 PI Donate F, Harris S, Plunkett ML, Mazar AP;
 XX WPI; 2002-666989/71.
 DR P-PSDB; ABN84910.
 XX
 PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
 PT diagnosing or treating diseases associated with undesired cell migration,
 PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
 XX
 PS Claim 1; Page 11; 82pp; English.
 XX
 CC The present sequence is the protein sequence of human histidine proline
 CC rich glycoprotein (HPRG), a proteinaceous metal chelator that can be used
 CC to inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
 CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
 CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
 CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
 CC endothelial cell proliferation or endothelial tube formation in vitro or
 CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
 CC Also claimed are: chemically synthesised or recombinantly produced
 CC peptide multimers; a diagnostically or therapeutically labeled anti-
 CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
 CC useful HPRG-related composition, comprising the diagnostically labeled
 CC polypeptide, peptide or peptide multimer and a carrier; an antibody
 CC specific for an epitope of HPRG that is present in the H/P domain of
 CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
 CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,
 CC or an antigen-binding fragment of the antibody; a method for inhibiting
 CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
 CC inducing apoptosis; a method for treating a subject having a disease or
 CC condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis; a method of stimulating or inhibiting
 CC angiogenesis in a subject; a method of detecting the presence of HPRG or
 CC its cleavage product or its peptide in a biological sample; isolated

Db	390	HPHGHHPHGHHPHGHHPHCHDFQDYGFCDPPH	422
		RESULT 6	
		ADH10409	
	ID	ADH10409 standard; protein; 525 AA.	
	XX		
	AC	ADH10409;	
	XX		
	DT	11-MAR-2004 (first entry)	
	XX		
	DE	Human HPRG protein sequence.	
	XX		
	KW	Tpm; tropomyosin; antiangiogenic receptor;	
	KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;	
	KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;	
	KW	antipsoriasis; dermatological; cardiac; vasotropic; vulnerary;	
	KW	angiogenesis; gene therapy; human.	
	XX		
	OS	Homo sapiens.	
	XX		
	Key	Location/Qualifiers	
	FH	276..321	
	FT	/note = pro-rich domain	
	FT	350..497	
	FT	/note = His-Pro rich domain	
	XX		
	PN	WO2003077872-A2.	
	XX		
	PD	25-SEP-2003.	
	XX		
	PF	17-MAR-2003; 2003WO-US008060.	
	XX		

CC	a biological sample; isolated nucleic acids encoding the polypeptide, peptide or peptide multimer; an expression vector; transformed or transfected cells; a method of providing to a cell, tissue or organ an angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer; an affinity ligand useful for binding to, or isolating, an HPRG-binding molecule or cells expressing the binding molecule, comprising the polypeptide, peptide or peptide multimer, immobilised to a solid support or carrier; and a method of isolating HPRG-binding molecule, or isolating or enriching cells expressing HPRG-binding site or receptor, from a complex mixture. The compositions and methods are useful in diagnosing or treating a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or fractures. HPRG is especially useful in inhibiting the growth of primary tumours or metastases, and may also be used in treating neurodegenerative diseases like Alzheimer's or Parkinson's disease. The antibodies are stimulators of angiogenesis and are useful for promoting neovascularization in pertinent disease states, and in various immunoassays	PS	Claim 6; SEQ ID NO 23; 117pp; English.
CC	Sequence 148 AA;	XX	The invention relates to an isolated tropomyosin (Tpm)-related antiangiogenic receptor polypeptide or peptide, which is a fragment of a full-length native Tpm protein expressed on the surface of endothelial cells, or a variant of the fragment. It has a molecular mass of about 17 kDa and corresponds in its sequence to, or is a variant of, an internal fragment of a native Tpm isoform which is a binding site for angiogenic polypeptide agents; The isolated antiangiogenic receptor polypeptide, peptide or variant has substantially the same biochemical activity of binding to the antiangiogenic polypeptide agents, as does the native Tpm internal fragment. The antiangiogenic polypeptide agent that binds to the isolated polypeptide or peptide is human histidine-proline rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic homologue, variant, domain or fragment of human or rabbit HPRG, two chain human kininogen (HK), the D5 domain of HK, or a Tpm-binding, antiangiogenic homologue, variant, domain or fragment of the HK or its D5 domain. The Tpm-related antiangiogenic receptor polypeptide or peptide, antibodies and compositions are useful for inhibiting endothelial cell migration, invasion, proliferation or angiogenesis, for inducing endothelial cell apoptosis, or for treating tumours or cancer, diabetic retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis, psoriasis, or scleroderma. The antibody may be also used for detecting the presence of a Tpm polypeptide or peptide in a biological sample, for promoting wound healing, or for treating diseases or conditions in which increased angiogenesis is desired, e.g. coronary artery disease or peripheral artery disease. The present sequence represents a human HPRG protein His-Pro (H/P) rich domain.
XX	Sequence 148 AA;	XX	Sequence 148 AA;
QY	Query Match 40.2%; Score 280.5; DB 5; Length 148; Best Local Similarity 51.6%; Pred. No. 2.3e-16; Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;	QY	Query Match 40.2%; Score 280.5; DB 8; Length 148; Best Local Similarity 51.6%; Pred. No. 2.3e-16; Matches 47; Conservative 1; Mismatches 24; Indels 19; Gaps 2;
QY	10 PHGHPHGGPPHGHHPHGGPPHGHHPHGGPPHGHHPHGGPPHGHHPHGGPPH 69	QY	10 PHGHPHGGPPHGHHPHGGPPHGHHPHGGPPHGHHPHGGPPHGHHPHGGPPH 69
DB	2 PHKHSHQHPHGHHPH-----HH-----PHEHDTTHROHPHGHHPHGHHP 42	DB	2 PHKHSHQHPHGHHPH-----HH-----PHEHDTTHROHPHGHHPHGHHP 42
QY	70 HGHPPHGGPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 100	QY	70 HGHPPHGGPPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 100
DB	43 HGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 73	DB	43 HGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPHGHHPH 73
RESULT 10		RESULT 11	
ID	ADH10411	ABG09949	
XX	ADH10411 standard; protein; 148 AA.	ID	ABG09949 standard; protein; 491 AA.
AC	ADH10411;	XX	ABG09949;
DT	11-MAR-2004 (first entry)	XX	13-FEB-2002 (first entry)
XX	Human HPRG protein H/P rich domain.	DE	Novel human diagnostic protein #9940.
XX	Tpm; tropomyosin; antiangiogenic receptor;	XX	Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;	KW	food supplement; medical imaging; diagnostic; genetic disorder.
KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;	XX	Homo sapiens.
KW	antipsoriatic; dermatological; cardiant; vasotropic; vulnary;	OS	WO200175067-A2.
XX	angiogenesis; gene therapy; human.	PN	11-OCT-2001.
OS	Homo sapiens.	PD	30-MAR-2001; 2001WO-US008631.
XX	WO2003077872-A2.	PF	31-MAR-2000; 2000US-00540217.
PN	25-SEP-2003.	PR	23-AUG-2000; 2000US-00649167.
XX	17-MAR-2003; 2003WO-US008060.	XX	(HYSE-) HYSEQ INC.
XX	15-MAR-2002; 2002US-0364047P.	PA	Drmanac RT, Liu C, Tang YT;
PR	(ATTE-) ATTENUON LLC.	XX	WPI; 2001-639362/73.
XX	McCrae K, Donate F, Juarez J, Mazar AP;		
PI	WPI; 2004-090604/09.		
XX	New tropomyosin-related antiangiogenic receptor polypeptide, useful for inhibiting endothelial cell migration, invasion, proliferation or angiogenesis, inducing endothelial cell apoptosis, or treating tumors or cancer.		

DR N-PSDB; AAS74136.
XX New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity.
XX
PS Claim 20; SEQ ID NO 40308; 103bp; English.
XX
XX The invention relates to isolated polynucleotide (I) and polypeptide (II)
CC sequences. (I) is useful as hybridisation probes, polymerase chain
CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping.
CC in and recombinant production of (II). The polynucleotides are also used
CC in diagnostics as expressed sequence tags for identifying expressed
CC genes. (I) is useful in gene therapy techniques to restore normal
CC activity of (II) or to treat disease states involving (II). (II) is
CC useful for generating antibodies against it, detecting or quantitating a
CC polypeptide in tissue, as molecular weight markers and as a food
CC supplement. (II) and its binding partners are useful in medical imaging
CC of sites expressing (II). (I) and (II) are useful for treating disorders
CC involving aberrant protein expression or biological activities. The
CC polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. ABG0010-ABG30377 represent novel human diagnostic
CC amino acid sequences of the invention. Note: The sequence data for this
CC patent did not appear in the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 491 AA;
XX
Query Match 37.9%; Score 264.5; DB 4; Length 491;
Best Local Similarity 41.8%; Pred. No. 1.3e-14;
Matches 56; Conservative 2; Mismatches 35; Indels 41; Gaps 8;
QY 6 HRPPP-----HGHPHG-----PPPHGHPH-GPPPHGHPGPPRPPHPP----- 45
Db 136 HRPPPAQVTTSTGHQHGHQHRPPPAAGHHQSSPPAAGHHQHRPPAAGHHQHGHQHS 195
QY 46 GPPPHGHPGPPPHGHPHGGPP-----PHGHPHGGPP-----HGH-----PPHG 86
Db 196 SPPAAGHHQHRHQPQHRSHPHSTAHHQHGHQSSPPAAGHHQHGHQHSPPYQ 255
QY 87 HGFPHDHCDCPPSH 100
Db 256 HGHHQHRP--PPAH 267
RESULT 12
ADD45442
ID ADD45442 standard; protein; 296 AA.
XX
AC ADD45442;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein L17318, SEQ ID NO 10875.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
OS Rattus norvegicus.
XX
PN WO2003016475-A2.
XX
PD 27-FEB-2003.
XX
PF 14-AUG-2002; 2002WO-US025765.
XX
PR 14-AUG-2001; 2001US-0312147P.
PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-0333347P.
XX
XX (GSHO) GEN HOSPITAL CORP.
PA (FARB) BAYER AG.
XX
PI Woolf C, D'urso D, Befort K, Costigan M;
XX
DR WPI; 2003-268312/26.
DR GENBANK; L17318.
XX
PT New composition comprising two or more isolated polypeptides, useful for
XX preparing a medicament for treating pain in an animal.
PS Claim 1; Page; 1017pp; English.
XX
CC The invention discloses a composition comprising two or more isolated rat
CC or human polynucleotides or a polynucleotide which represents a fragment,
CC derivative or allelic variation of the nucleic acid sequence. Also
CC claimed are a vector comprising the novel polynucleotide, a host cell
CC comprising the vector, a method for identifying a nucleotide sequence
CC which is differentially regulated in an animal subjected to pain and a
CC kit to perform the method, an array, a method for identifying an agent
CC that increases or decreases the expression of the polynucleotide sequence
CC that is differentially expressed in neuronal tissue of a first animal
CC subjected to pain, a method for identifying a compound which regulates
CC the expression of a polynucleotide sequence which is differentially
CC expressed in an animal subjected to pain, a method for identifying a
CC compound that regulates the activity of one or more of the
CC polynucleotides, a method for producing a pharmaceutical composition, a
CC method for identifying a compound or small molecule that regulates the
CC activity in an animal of one or more of the polypeptides given in the
CC specification, a method for identifying a compound useful in treating
CC pain and a pharmaceutical composition comprising the one or more
CC polypeptides or their antibodies. The polynucleotide or the compound that
CC modulates its activity is useful for preparing a medicament for treating
CC pain (e.g. spinal segmental nerve injury (SNI)) in an animal (e.g. gene
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
CC therapy). The sequence presented is a rat protein (shown in Table 2 of
CC the specification) which is differentially expressed during pain. Note:
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic form directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 296 AA;
XX
Query Match 37.2%; Score 259.5; DB 7; Length 296;
Best Local Similarity 38.4%; Pred. No. 2.3e-14;
Matches 63; Conservative 1; Mismatches 27; Indels 73; Gaps 10;
QY 2 VNIITHRPPPHG---HHPH-----GPPPHG---HHPHGPPPHGHP-- 34
Db 81 VNRPERPPQHGNGHHHPHPPAAGPQRPQPGSPQGPQRPQRPQGPQGPQRPQ 140
QY 35 -----PHGPPP-----RHPPHGGPPPHGHP-----PHGPPPHGHP----PHGPPPHG 71
Db 141 PPQFGSPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQ 200
QY 72 HP-----PHGP-----PPHGHPPHGHGHDHGDGCDPP 98
Db 201 GPQRPQPGSPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGPQGP 238
RESULT 13
ADE57105
ID ADE57105 standard; protein; 296 AA.
XX
AC ADE57105;
XX
DT 29-JAN-2004 (first entry)
XX
DE Rat Protein L17318, SEQ ID NO 2965.
XX
KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;

Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 Rattus norvegicus.
 WO2003016475-A2.
 27-FEB-2003.
 14-AUG-2002; 2002WO-US025765.
 14-AUG-2001; 2001US-0312147P.
 01-NOV-2001; 2001US-0346382P.
 26-NOV-2001; 2001US-0333347P.
 (GEO) GEN HOSPITAL CORP.
 (FARB) BAYER AG.
 Woolf C, D'urso D, Befort K, Costigan M;
 WPI; 2003-268312/26.
 GENBANK; L17318.
 New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
 Claim 1; Page; 1017pp; English.
 The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (SNI)), in an animal (e.g. gene injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 Sequence 296 AA;
 Query Match 37.2%; Score 259.5; DB 7; Length 296;
 Best Local Similarity 38.4%; Pred. No. 2.3e-14;
 Matches 63; Conservative 1; Mismatches 27; Indels 73; Gaps 10;
 QY 2 VNIHRRPPHG---HHPH-----GPPPHG---HHPHGPPPHGHP-- 34
 Db 81 VNRPERPQGGNHHPHPPAAGPQRPQPGSPQGGPQRRPQGGPQGPQR 140
 QY 35 -----PHGPP-----RHPHGPHPHGP-----PHGPPHGP-----PHGPPHGP 71
 Db 141 PPQFGSPQGGPQGPQRRPQGGPQRRPQGGPQRRPQGGPQRRPQGGPQ 200
 QY 72 HP-----PHGP-----PPHGHPPHGHGHDGCDPP 98
 Db 201 GPORPQPGSPQGGPQRRPQGGPQRRPQGGPQRRPQGGPQRRPQGGPQ 238

RESULT 14
 ADD47240
 ID ADD47240 standard; protein; 274 AA.
 XX
 AC ADD47240;
 XX
 DT 02-DEC-2004 (revised)
 DT 29-JAN-2004 (first entry)
 XX
 DE Rat Protein AAA42064, SEQ ID NO 12934.
 XX
 KW Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW Chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX
 OS Rattus norvegicus.
 OS Unidentified.
 XX
 PN WO2003016475-A2.
 XX
 PD 27-FEB-2003.
 XX
 PF 14-AUG-2002; 2002WO-US025765.
 XX
 PR 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-0333347P.
 XX
 PA (GEO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.
 XX
 PI Woolf C, D'urso D, Befort K, Costigan M;
 XX
 WPI; 2003-268312/26.
 GENBANK; AAA42064.
 XX
 PT New composition comprising two or more isolated polypeptides, useful for preparing a medicament for treating pain in an animal.
 XX
 PS Example 1; Page; 1017pp; English.
 XX
 CC The invention discloses a composition comprising two or more isolated rat or human polynucleotides or a polynucleotide which represents a fragment, derivative or allelic variation of the nucleic acid sequence. Also claimed are a vector comprising the novel polynucleotide, a host cell comprising the vector, a method for identifying a nucleotide sequence which is differentially regulated in an animal subjected to pain and a kit to perform the method, an array, a method for identifying an agent that increases or decreases the expression of the polynucleotide sequence that is differentially expressed in neuronal tissue of a first animal subjected to pain, a method for identifying a compound which regulates the expression of a polynucleotide sequence which is differentially expressed in an animal subjected to pain, a method for identifying a compound that regulates the activity of one or more of the polynucleotides, a method for producing a pharmaceutical composition, a method for identifying a compound or small molecule that regulates the activity in an animal of one or more of the polypeptides given in the specification, a method for identifying a compound useful in treating pain and a pharmaceutical composition comprising the one or more polypeptides or their antibodies. The polynucleotide or the compound that modulates its activity is useful for preparing a medicament for treating pain (e.g. spinal segmental nerve injury (SNI)), in an animal (e.g. gene injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene therapy). The sequence presented is a rat protein (shown in Table 2 of the specification) which is differentially expressed during pain. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic form directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
 CC
 CC Sequence 274 AA;
 Query Match 35.3%; Score 246; DB 7; Length 274;
 Best Local Similarity 47.1%; Pred. No. 2.9e-13;

Result No.	Score	Query Match	Length	DB	ID	Description	
1	314	45.1	525	4	US-09-976-594-64	Sequence 64, Appl	
2	314	45.1	525	4	US-09-919-039-62	Sequence 62, Appl	
3	233.5	33.5	321	4	US-09-538-092-845	Sequence 845, App	
4	233.5	33.5	424	4	US-09-538-092-1338	Sequence 1338, Ap	
5	231	32.9	320	4	US-09-949-016-8196	Sequence 8196, Ap	
6	229.5	32.9	124	2	US-08-925-237-2	Sequence 2, Appl	
7	229.5	32.9	251	4	US-09-538-092-840	Sequence 840, App	
8	228	32.7	247	4	US-09-538-092-890	Sequence 890, App	
9	224	32.1	276	4	US-09-538-092-889	Sequence 889, App	
10	221.5	31.8	234	4	US-09-538-092-888	Sequence 888, App	
11	220	31.6	258	4	US-09-949-016-9409	Sequence 9409, Ap	
12	215.5	30.9	104	4	US-09-547-693-235	Sequence 235, App	
13	208	29.8	865	3	US-09-281-766-19	Sequence 19, Appl	
14	208	29.8	865	4	US-09-612-858-19	Sequence 19, Appl	
15	208	29.8	865	4	US-09-957-995A-19	Sequence 19, Appl	
16	208	29.8	971	4	US-09-248-796A-19531	Sequence 19, Appl	
17	205	29.4	130	4	US-09-270-767-39206	Sequence 39206, A	
18	205	29.4	130	4	US-09-270-767-5423	Sequence 5423, A	
19	196	28.1	82	2	US-08-925-237-4	Sequence 4, Appl	
20	192.5	27.6	511	4	US-09-107-433-2904	Sequence 2904, Ap	
21	192.5	27.6	955	4	US-09-949-016-8369	Sequence 8369, Ap	
22	192	27.5	324	4	US-09-949-016-7664	Sequence 7664, Ap	
23	191.5	27.5	553	4	US-09-949-016-7961	Sequence 7961, Ap	
24	190	27.3	166	4	US-09-949-016-6709	Sequence 6709, Ap	
25	190	27.3	178	4	US-09-949-016-8394	Sequence 8394, Ap	
26	175	25.1	306	1	US-08-217-327-6	Sequence 6, Appl	
27	175	25.1	623	4	US-09-949-016-6530	Sequence 6530, Ap	

```
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 085596CD1
US-09-919-039-62

Query Match      45.1%; Score 314; DB 4; Length 525;
Best Local Similarity 41.8%; Pred. No. 4.6e-21;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

QY 7 RPP--PHG-----HHPGPPPHGHHHPGPP-----HGHP-PHGPPP-----40
Db 275 KPPFKGHSRDHHHPHKP-----HEHGPPPPDPRDHSHPPLPQGGPPLPMSSCSQH 329
QY 41 -----RHP-----PHGPPPHGHP-----HGPPPHGHPGCP 67
Db 330 ATFGTGAQRSSNNSSDLPHKHKSHSQHPGHHPHHAHPHEHDTHRQHPGHHPGH 389
QY 68 PPHGHPHPGPPPHGHPHGHGFDHGFCDPPSH 100
Db 390 HPHGHPHGHHPHGHHPCHDFDYGCDPPPH 422

RESULT 3
US-09-538-092-845
; Sequence 845, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 845
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P04280
US-09-538-092-845

Query Match      33.5%; Score 233.5; DB 4; Length 331;
Best Local Similarity 42.2%; Pred. No. 5.1e-14;
Matches 49; Conservative 4; Mismatches 38; Indels 25; Gaps 5;

QY 8 PPHGHHHPGPPPHGHHHPGPPH-GHPHGGPPPHHPHGGPPPHG-----PHHG 56
Db 101 PPGGNGPQGPQPPGKPGQPPGKPGQPPGKPGQPPGKPGQPPGKPGQPPGKPGQ 160
QY 57 PPHH-GHPHGGPP-----HGHPHGGPPHGH-PHGHPHGHGHPGCDPP 98
Db 161 PPGGNGPQGPQPPGKPGQPPGKPGQPPGKPGQPPGKPGQPPGKPGQPPGKPG 216

RESULT 4
US-09-538-092-1338
; Sequence 1338, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CuraPatSeqFormatter Version 0.9
; SEQ ID NO 1338
; LENGTH: 424
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number Q15427
US-09-538-092-1338

Query Match      33.5%; Score 233.5; DB 4; Length 424;
Best Local Similarity 46.0%; Pred. No. 6.4e-14;
Matches 57; Conservative 3; Mismatches 25; Indels 39; Gaps 11;

QY 9 PPHGH-HPHG-PP-----PHG-HPHGPP-----PHG-HPHGPP-----HGHPHGGPPPHG 46
Db 286 PCHGSHHPHPPFPGGPHPGMSQMLAHGHPHGLHHPHGGGQPPPPGMPHG 345
QY 47 PPHGHPHGP-----PHGHPHGP-PHGH-PHHPHGP-----PPHGHGPHDHP 94
Db 346 PPMGMPHPPGPPFGSPMGP--GPMPPHGMGPPPLMPHGYTGPPPPPYG---YQKGP 400
QY 95 CDPP 98
Db 401 LPPP 404

RESULT 5
US-09-949-016-8196
; Sequence 8196, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8196
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8196

Query Match      33.1%; Score 231; DB 4; Length 320;
Best Local Similarity 38.6%; Pred. No. 8.4e-14;
Matches 51; Conservative 6; Mismatches 37; Indels 38; Gaps 6;

QY 8 PPHGHHHPGPPPHGH-----PHGPPH-GHPHGGPPR-HPHGPPPH-----50
```

Db 166 PPHPGKPEGPPQGGNQSQGPPRPGKPEGPPPPQGGNQSQGPPRPGKPEGPPPPQGGNQ 225
 Qy 51 -----GHPPHGGPPRHHPPHGGPPR-HCHPHGPPRHHGHP--PHGHGF 89
 Db 226 SQGPPRPGKPEGPPSQGGNKGPPPPHFGKPPGPPPPQGGNKGPPRPPRPGKPEGPPPPG 285
 Qy 90 HDHGPCDPPSHK 101
 Db 286 NPOQLPPPPACK 297

RESULT 6
 US-08-925-237-2
 ; Sequence 2, Application US/08925237
 ; Patent No. 5981720
 ; GENERAL INFORMATION:
 ; APPLICANT: Azen, Edwin A.
 ; APPLICANT: Pan, David
 ; TITLE OF INVENTION: Human Salivary Proteins And Fragments
 ; TITLE OF INVENTION: Thereof Having Alpha-Glucosidase Inhibitory Activity
 ; NUMBER OF SEQUENCES: 5
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Kent Barta
 ; STREET: 100 East Wisconsin Avenue
 ; CITY: Milwaukee
 ; STATE: WI
 ; COUNTRY: USA
 ; ZIP: 53202-4108
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/925,237
 ; FILING DATE:
 ; CLASSIFICATION: 514
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Barta, Kent S.
 ; REGISTRATION NUMBER: 29,042
 ; REFERENCE/DOCKET NUMBER: 96429/9002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (608) 257-3501
 ; TELEFAX: (608) 283-2275
 ; INFORMATION FOR SEQ ID NO: 2:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 124 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-925-237-2

Query Match 32.9%; Score 229.5; DB 2; Length 124;
 Best Local Similarity 44.9%; Pred. No. 4.8e-14;
 Matches 48; Conservative 3; Mismatches 35; Indels 21; Gaps 5;
 Qy 8 PPHGHHPHGGPPPHGHPHPPH-GHPHGGPPRHHPPHGGPPHGH-----PPH 55
 Db 10 PQGGNQPPGPPPPGKPPGPPPPQGGNKGPPPPGKPPGPPPPQGGNKSARSPPGKQ 69
 Qy 56 GPPPH-GHPHGGPPHGGPPHGGPPHGGHGHGHDHGPCDPPSHK 101
 Db 70 GPPPGNQPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGK 108

RESULT 7
 US-09-538-092-840
 ; Sequence 840, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same

FILE REFERENCE: 15966-542
 CURRENT APPLICATION NUMBER: US/09/538,092
 CURRENT FILING DATE: 2000-03-29
 PRIOR APPLICATION NUMBER: 60/127,352
 PRIOR FILING DATE: 1999-04-01
 PRIOR APPLICATION NUMBER: 60/178,965
 PRIOR FILING DATE: 2000-02-01
 NUMBER OF SEQ ID NOS: 1387
 SOFTWARE: CuratSeqFormatter Version 0.9
 SEQ ID NO 840
 LENGTH: 251
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (0)...(0)
 OTHER INFORMATION: Polypeptide Accession Number P02812
 US-09-538-092-840

Query Match 32.9%; Score 229.5; DB 4; Length 251;
 Best Local Similarity 43.6%; Pred. No. 9.1e-14;
 Matches 48; Conservative 3; Mismatches 34; Indels 25; Gaps 5;
 Qy 8 PPHGHHPHGGPPPHGHPHPPH-GHPHGGPPRHHPPHGGPPHGH-----PPH 55
 Db 19 PQGGNQPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGK 78
 Qy 56 GPPPH-GHPHGGPPHGGPPHGGPPHGGH-----PPHGHGHDHGPCDPP 98
 Db 79 GPPPGNQPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGKPPGPPPPGK 122

RESULT 8
 US-09-538-092-890
 ; Sequence 890, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Giot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 SOFTWARE: CuratSeqFormatter Version 0.9
 SEQ ID NO 890
 LENGTH: 247
 TYPE: PRT
 ORGANISM: Homo sapiens
 FEATURE:
 NAME/KEY: misc_feature
 LOCATION: (0)...(0)
 OTHER INFORMATION: Polypeptide Accession Number P10163
 US-09-538-092-890

Query Match 32.7%; Score 228; DB 4; Length 247;
 Best Local Similarity 36.4%; Pred. No. 1.2e-13;
 Matches 52; Conservative 4; Mismatches 39; Indels 48; Gaps 5;
 Qy 7 RPPHGHHPHGGPPPHGHPHPPH-GHPHGGPPRHHPPHGGPPHGH-----PPHGH 52
 Db 81 RPPQGGNQSQGPPPPHGGKPPRPPGPPGPPGPPGPPGPPGPPGPPGPPGPPG 140
 Qy 53 P-----PHGPPPHGHPHGGPP-----HGHPHGGPPH 80
 Db 141 PERPPPGNQSQGPPPPHGGKPPGPPPPGPPGPPGPPGPPGPPGPPGPPGPP 200
 Qy 81 GHP--PHGHPHGHGCDPPSHK 101

```

Db      201 GKPQGPAGGNPQQPDPAGK 223

RESULT 9
US-09-538-092-889
; Sequence 889, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 889
; LENGTH: 276
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P10162
US-09-538-092-889

Query Match      32.1%; Score 224; DB 4; Length 276;
Best Local Similarity 37.6%; Pred. No. 3.1e-13;
Matches 50; Conservative 3; Mismatches 36; Indels 44; Gaps 5;

QY      7 RPPHGHHPGPPHGHHPGPPH-----HPPHGGPPHPP-----HGPPRHGP 53
DB      131 RPPQGNQSQGPPHPGKPERPPPPQGGNQSRRPPPPGKPERPPPPQGGNQSRRPP 190
QY      54 PHGPP-----HGPPHGGPPHGHHPGPPHGH-----PPH 85
DB      191 PEGPPQEGNKSARSPPGKQGPQQEGNKFQGPFPFGKQGPFPFGNPGNQQA 250
QY      86 GHGFHDHGCDPP 98
DB      251 G---KPQGP 260

RESULT 10
US-09-538-092-888
; Sequence 888, Application US/09538092
; Patent No. 6753314
; GENERAL INFORMATION:
; APPLICANT: Giot, Loic
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqFormatter Version 0.9
; SEQ ID NO 888
; LENGTH: 234
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P10161
US-09-538-092-888

Query Match      31.8%; Score 221.5; DB 4; Length 234;
Best Local Similarity 42.3%; Pred. No. 4.5e-13;
Matches 47; Conservative 5; Mismatches 40; Indels 19; Gaps 5;

QY      7 RPPHGHHPGPPH-GHHPHGPPPHGHPHGGPPR--HPPHGGPPH-----GH 52
DB      15 RPPPPGKPGQPPPGQGNQSQGPPPHGKPERPPPPQGGNQSRRPPPGKPERPPPGGN 74
QY      53 PPHGPPPHGHPHGGPPH-GHPPHGGPPH---GHPPHGHGHGHDGCDPP 98
DB      75 QSQGPPTPGKPEGPPPGQGNQSQGPPPHGKPERPPPPQGGNQSRRPPPPP 125

RESULT 11
US-09-949-016-9409
; Sequence 9409, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9409
; LENGTH: 258
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9409

Query Match      31.6%; Score 220; DB 4; Length 258;
Best Local Similarity 35.7%; Pred. No. 6.7e-13;
Matches 51; Conservative 4; Mismatches 40; Indels 48; Gaps 5;

QY      7 RPPHGHHPGPPHGHHPHGGPPH-----HPPHGGPPPHPPHG-----PPPHGH 52
DB      92 RPPQGNQSQGPPPHGKPERPPPPQGGNQSQTTPPPGKPERPPPGNQSRRPPPGK 151
QY      53 P-----PHGPPPHGHPHGGPP-----HGHPHGGPPH 80
DB      152 PERPPPGQGNQSQGPPPHGKPEGPPPGQEGNKSARSPPGKQGPQQEGNKGKPGPPPP 211
QY      81 GHP--PHGHPHDHGCDPPSHK 101
DB      212 GKQGPAGNPGNPGQQA 234

RESULT 12
US-09-547-693-235
; Sequence 235, Application US/09547693
; Patent No. 6639050
; GENERAL INFORMATION:
; APPLICANT: Kieliszewski, Marcia
; TITLE OF INVENTION: Synthetic Genes for Plant Gums and Other Hydroxyproline-Rich
; TITLE OF INVENTION: Glycoproteins
; FILE REFERENCE: OHU-04089
; CURRENT APPLICATION NUMBER: US/09/547,693
; CURRENT FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 236
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 235
; LENGTH: 104
; TYPE: PRT
; ORGANISM: Artificial/Unknown

```

```

; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Synthetic
US-09-547-693-235

```

[illegible]

```

RESULT 13
US-09-281-766-19
; Sequence 19, Application US/09281766
; Patent No. 6376196
; GENERAL INFORMATION:
; APPLICANT: Conrad, Patricia C.
; APPLICANT: Louie, Kitland
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Recombinant Neosporea Antigens and Their Uses
; FILE REFERENCE: 023070-082510US
; CURRENT APPLICATION NUMBER: US/09/281,766
; CURRENT FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 08/645,951
; PRIOR FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 19
; LENGTH: 865
; TYPE: PR1
; ORGANISM: Neosporea sp.
; OTHER INFORMATION: NC-p65 cDNA
US-09-281-766-19

```

```

Query Match      29.8%; Score 208; DB 3; Length 865;
Best Local Similarity 37.9%; Pred. No. 2.4e-11;
Matches 39; Conservative 5; Mismatches 47; Indels 12; Gaps 2;

QY      8 PPRHGHHGPPRRGHGHHGPPRRGHGHPHPPRR-----HPRHGPPRHGHPRHG 56
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      559 PPEASPPSSPDRQHPRHPRPRNPPEASPPSPSPNNQWHPHPRPRNPPEASPP 618

QY      57 PPRHGHPHPPRRGHGHPHPPRRGHGHPHPPRHGHGPHDHGP-CDDP 98
      ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      619 SPNNQWHPHPRPRNPPEASPPSPSPNNQWHPHPRPRNP 661
      ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 14
US-09-612-858-19
; Sequence 19, Application US/09612858
; Patent No. 6716423
; GENERAL INFORMATION:
; APPLICANT: Conrad, Patricia C.
; APPLICANT: Louie, Kitland
; APPLICANT: The Regents of the University of California
; TITLE OF INVENTION: Recombinant Neospora Antigens and Their Uses
; FILE REFERENCE: 023070-082510US
; CURRENT APPLICATION NUMBER: US/09/612,858
; CURRENT FILING DATE: 2000-07-10
; PRIOR APPLICATION NUMBER: US/09/281,766
; PRIOR FILING DATE: 1999-03-30
; PRIOR APPLICATION NUMBER: US 08/645,951
; PRIOR FILING DATE: 1996-05-10
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 19

```

; LENGTH: 865
; TYPE: PRT
; ORGANISM: Neospora sp.
; OTHER INFORMATION: NC-p65 cDNA
US-09-612-858-19

```

```

Query Match      29.8%; Score 208; DB 4; Length 865;
Best Local Similarity 37.3%; Pred. No. 2.4e-11;
Matches 39; Conservative 5; Mismatches 47; Indels 12; Gaps 2;

Qy      8 PPRGHNHPPRPHGHNHPPRPHGHPPR-----HPPGPPPHGHPPHG 56
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      559 PPEASPPSSPDROQHPPRPPRPPASPPSPPPNQHPPRPPRPPASPP 618

Qy      57 PPRGHPPRPPGHPPRPPGHPPRPPGHPPGHGPHDHP-CDPP 98
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      619 SPNWOHPPHPPRPPRPPASPPSSPPSPNWOHPPRPPRPP 661
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-957-995A-19
; Sequence 19, Application US/09957995A
; Patent No. 6777192
; GENERAL INFORMATION:
; APPLICANT: Conrad, Patricia A.
; Barr, Bradd C.
; Anderson, Mark L.
; Sverlow, Karen W.
; TITLE OF INVENTION: Recombinant Neospora Antigens and Their
; Uses
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk

```

```

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/09/957,995A
  FILING DATE: 15-Apr-2002
  CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
  APPLICATION NUMBER: US 08/215,858
  FILING DATE: 21-MAR-1994
  APPLICATION NUMBER: US 08/327,516
  FILING DATE: 20-OCT-1994
  APPLICATION NUMBER: US 08/645,951
  FILING DATE: 10-MAY-1996
  APPLICATION NUMBER: US 09/281,766
  FILING DATE: 30-MAR-1999
ATTORNEY/AGENT INFORMATION:
  NAME: Bastian, Kevin L.
  REGISTRATION NUMBER: 34,774
  REFERENCE/DOCKET NUMBER: 023070-082512US
TELECOMMUNICATION INFORMATION:
  TELEPHONE: (415) 576-0200
  TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 19:
  SEQUENCE CHARACTERISTICS:
    LENGTH: 865 amino acids
    TYPE: amino acid
    TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 19:
US-09-957-995A-19

```

Query Match 29.8%; Score 208; DB 4; Length 865;

0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:37:20 ; Search time 19.4699 seconds
(without alignments)
499.124 Million cell updates/sec

Title: US-10-074-225a-6
Perfect score: 697
Sequence: 1 SVNIHRPPPHGHHPGPPPP.....HPPHGHGHGDCDPPPSHK 101

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: PIR1:.*
2: PIR2:.*
3: PIR3:.*
4: PIR4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	314	45.1	525	1 KGHUGH	histidine-rich gly
2	260.5	37.4	445	2 A60488	histidine-rich gly
3	259.5	37.2	295	2 B48013	proline-rich prote
4	250	35.9	128	2 D38355	basic proline-rich
5	248	35.6	278	2 B39066	proline-rich prote
6	247.5	35.5	206	1 PIR73	acidic proline-ric
7	246.5	35.4	227	2 C29149	proline-rich prote
8	246.5	35.4	240	2 A24264	proline-rich prote
9	246.5	35.4	317	2 A28996	proline-rich prote
10	239	34.3	240	2 B24264	proline-rich prote
11	236	33.9	330	2 S22140	nodulin Enod2 - Se
12	235.5	33.8	392	1 PIHUB6	salivary proline-r
13	235.5	33.8	432	2 T06782	extensin - soybean
14	235.5	33.8	580	2 T10863	extensin precursor
15	234.5	33.6	173	2 T51469	glycine/proline-ri
16	233.5	33.5	309	2 S08343	nodulin precursor
17	233.5	33.5	424	2 A54964	spliceosome-associ
18	232.5	33.4	1571	2 T14155	zinc finger protei
19	229.5	32.9	251	1 PIHUFF	salivary proline-r
20	228.5	32.8	188	2 D29149	proline-rich prote
21	225.5	32.4	212	2 B36298	proline-rich prote
22	224	32.1	310	1 PIHUSD	salivary proline-r
23	222.5	31.9	220	2 A36298	proline-rich prote
24	221	31.7	309	2 S10889	proline-rich prote
25	221	31.7	1378	2 T30173	zinc finger protei
26	214.5	30.8	117	2 D40750	proline-rich prote
27	213.5	30.6	301	2 E29149	proline-rich prote
28	210.5	30.2	434	2 S11967	nodule-specific by
29	210	30.1	297	2 S23737	proline-rich prote

30	210	30.1	440	1 S60755	rhodopsin - Allote
31	210	30.1	620	2 S06733	hydroxyproline-ric
32	209	30.0	97	2 S04554	nodulin - alfalfa
33	208	29.8	223	2 A42817	proline-rich prote
34	207	29.7	727	2 C84534	hypothetical prote
35	206.5	29.6	112	2 S10101	modulation protein
36	206.5	29.3	539	2 T28770	hypothetical prote
37	202.5	29.1	170	2 A48013	proline-rich prote
38	202	29.0	168	2 JH0481	basic proline-rich
39	201	28.8	452	1 S14332	rhodopsin - northe
40	199.5	28.6	172	2 B29149	proline-rich prote
41	198.5	28.5	1176	2 T49482	hypothetical prote
42	198	28.4	164	2 A30496	proline-rich prote
43	197.5	28.3	260	2 S22373	proline-rich prote
44	197.5	28.3	592	2 T32402	hypothetical prote
45	197.5	28.3	1002	2 T43236	carboxypeptidase C

ALIGNMENTS

RESULT 1

KGHUGH
histidine-rich glycoprotein precursor - human
N:Alternate names: HRG
C:Species: Homo sapiens (man)
C>Date: 04-Dec-1986 #sequence_revision 04-Dec-1986 #text_change 09-Jul-2004
C:Accession: A01287; S29669
R:Koide, T.; Foster, D.; Yoshitake, S.; Davie, E.W.
Biochemistry 25, 2220-2225, 1986
A:Title: Amino acid sequence of human histidine-rich glycoprotein derived from the nucleolus
A:Reference number: A01287; MUID:86216149; PMID:3011081
A:Accession: A01287
A:Molecule type: mRNA
A:Residues: 1-525 <KOI>
A:Cross-references: UNIPROT:P04196; GB:AB005803; NID:g2280513; PIDN:BAA21613.1; PID:g2280513
R:Henniss, B.; Havelaar, A.; Klufft, C.
Submitted to the EMBL Data Library, October, 1991
A:Description: PCR detection of a dinucleotide repeat in the human histidine-rich glycoprotein
A:Reference number: S29669
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 214-247 <HEN>
A:Cross-references: EMBL:217218; NID:g32453; PIDN:CAA78925.1; PID:g32454
C:Comment: Although its physiological function is not yet known, HRG does bind heme, dyef
din, and the lysine-binding site of plasminogen. On the basis of its homology with HMW k
load coagulation cascade.
C:Comment: The amino half of this protein is homologous to the first two cystatin-like d
ould not have inhibitory activity.
C:Comment: In addition to having a high histidine and proline content, this protein has n
e-rich' region.
C:Genetics:
A:Gene: GDB:HRG
A:Cross-references: GDB:120055; OMIM:142640
A:Map position: 3q27-3q27
C:Superfamily: histidine-rich glycoprotein; cystatin homology
C:Keywords: duplication; glycoprotein; heparin binding; tandem repeat
F:1-18/Domain: signal sequence #status predicted <SIG>
F:19-525/Product: histidine-rich glycoprotein #status predicted <MAT>
F:19-131/Domain: cystatin homology <CV1>
F:140-246/Domain: cystatin homology <CY2>
F:276-321/Region: proline-rich
F:348-437/Region: histidine-rich
F:351-497/Region: proline-rich
F:63.125.344.345/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:78-89.105-126.218-241/Disulfide bonds: #status predicted

Query Match 45.1% Score 314; DB 1; Length 525;
Best Local Similarity 41.8%; Pred. No. 1.7e-16;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;
QY 7 RPP--PHG-----HHPHGPPPHGHHPGPPPP-----HGHP-PHGPPP----- 40

A:Cross-references: UNIPROT:Q04154; GB:M36414
C:Superfamily: proline-rich protein

Query Match 35.6%; Score 248; DB 2; Length 278;
Best Local Similarity 47.2%; Pred. No. 6.7e-12;
Matches 58; Conservative 1; Mismatches 32; Indels 32; Gaps 9;

QY 8 RPPHGHHP-----HHPHGPPPHG-----HHPHGPPPHG 51
DB 52 PPPKGGQQRRPPPGGQQKPPGPPPGGQQRRPPPGGQQRRPPPG 111
QY 52 HP---PHGPPPHGHP---PHGPPPHGHP---PHGPPPHG---HPPHG---HGFDHGPC 95
DB 112 GPQRRPQGGPPPGGPPPGGPPPGGPPPGGPPPGGPPPGGPPG 171
QY 96 DPP 98
DB 172 GPP 174

RESULT 6

PIRT3
acidic proline-rich protein precursor - rat
N:Alternate names: PRP
C:Species: Rattus norvegicus (Norway rat)
C:Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A03296
R:Ziemer, M.A.; Swain, W.F.; Rutter, W.J.; Clements, S.; Ann, D.K.; Carlson, D.M.
J. Biol. Chem. 259, 10475-10480, 1984
A:Title: Nucleotide sequence analysis of a proline-rich protein cDNA and peptide homolog
A:Reference number: A03296; MUID:84289443; PMID:6547951
A:Accession: A03296
A:Molecule type: mRNA
A:Residues: 1-206 <ZIE>
A:Cross-references: UNIPROT:P04474; GB:K02247; NID:G206395; PIDN:AAA41949.1; PID:G206396
C:Comment: This protein contains six 18- to 19-residue repeats.
C:Comment: This protein may protect teeth by binding to tannins.
C:Superfamily: proline-rich protein
C:Keywords: duplication; parotid gland; saliva; tandem repeat
F:1-13/Domain: signal sequence #status predicted <SIG>
F:14-206/Product: acidic proline-rich protein #status predicted <MAT>
F:80-189/Region: 18-residue repeats

Query Match 35.5%; Score 247.5; DB 1; Length 206;
Best Local Similarity 48.2%; Pred. No. 5.6e-12;
Matches 53; Conservative 4; Mismatches 30; Indels 23; Gaps 7;

QY 7 RPPHGHHPGPPPHGHPGPPPHGHP-----PHGPPPHGPPPHGPPPHGPPPHGPP 58
DB 79 RPPPHGNNHRRPP-GHH-HGPPPSGGPQTSSQGNPQGNPQGGPQGGPQGN-PQGGP 135
QY 59 PHGHP-----PHGPPPHGHPGPPPHGHP--PHGHPGFDHGPCDP 97
DB 136 PQGGPQRRPPGKQPGPPPGG-GPQGPQGNPQGNPQGNPQGGHQQRRPPQ 184

RESULT 7

C29149
proline-rich protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 08-Mar-1989 #sequence_revision 08-Mar-1989 #text_change 09-Jul-2004
C:Accession: C29149
R:Clements, S.; Mehanho, H.; Carlson, D.M.
J. Biol. Chem. 260, 13471-13477, 1985
A:Title: Novel multigene families encoding highly repetitive peptide sequences. Sequence
A:Reference number: A92501; MUID:86033759; PMID:3840480
A:Contents: Clone pUMP40
A:Accession: C29149
A:Molecule type: mRNA
A:Residues: 1-227 <CLE>
A:Cross-references: UNIPROT:Q62107; GB:M11902; NID:G200544; PIDN:AAA40003.1; PID:G200545
C:Superfamily: proline-rich protein

Query Match 35.4%; Score 246.5; DB 2; Length 227;
Best Local Similarity 43.6%; Pred. No. 7.2e-12;
Matches 58; Conservative 0; Mismatches 34; Indels 41; Gaps 9;

QY 7 RPPPHGHHP-----HGPPPHG-----HHPHGPPPHG-----HHPHGPPPHG 45
DB 33 RPPPSGQQRRPPVNGSQGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 92
QY 46 GPPPHG-----HHPHGPPPHG-----HHPHGPPPHG-----HHPHGPPPHG--- 86
DB 93 GPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 152
QY 87 -HGFDHGPCDPP 98
DB 153 PAGQPPRPPGPP 165

RESULT 8

A24264
proline-rich protein MP2 - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 09-Sep-1987 #sequence_revision 09-Sep-1987 #text_change 03-May-1996
C:Accession: A24264
R:Ann, D.K.; Carlson, D.M.
J. Biol. Chem. 260, 15863-15872, 1985
A:Title: The structure and organization of a proline-rich protein gene of a mouse multigene
A:Reference number: A92508; MUID:86059475; PMID:2999141
A:Accession: A24264
A:Molecule type: DNA
A:Residues: 1-240 <ANN>
C:Superfamily: proline-rich protein

Query Match 35.4%; Score 246.5; DB 2; Length 240;
Best Local Similarity 43.6%; Pred. No. 7.6e-12;
Matches 58; Conservative 0; Mismatches 34; Indels 41; Gaps 9;

QY 7 RPPPHGHHP-----HGPPPHG-----HHPHGPPPHG-----HHPHGPPPHG 45
DB 12 RPPPSGQQRRPPVNGSQGPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 71
QY 46 GPPPHG-----HHPHGPPPHG-----HHPHGPPPHG-----HHPHGPPPHG--- 86
DB 72 GPPPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPPGPP 131
QY 87 -HGFDHGPCDPP 98
DB 132 PAGQPPRPPGPP 144

RESULT 9

A28996
proline-rich protein M14 precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change 09-Jul-2004
C:Accession: A28996
R:Ann, D.K.; Smith, K.; Carlson, D.M.
J. Biol. Chem. 263, 10887-10893, 1988
A:Title: Molecular evolution of the mouse proline-rich protein multigene family. Insertion
A:Reference number: A28996; MUID:88273214; PMID:2839509
A:Accession: A28996
A:Molecule type: DNA
A:Residues: 1-317 <ANN>
A:Cross-references: UNIPROT:Q62103; GB:M23236; GB:J03891; NID:G200535; PIDN:AAA53048.1; PID:G200536
C:Genetics:
A:Introns: 22/1
C:Superfamily: proline-rich protein
C:Keywords: saliva
F:1-15/Domain: signal sequence #status predicted <SIG>
F:16-317/Product: proline-rich protein M14 #status predicted <MAT>

Query Match 35.4%; Score 246.5; DB 2; Length 317;
Best Local Similarity 43.6%; Pred. No. 9.6e-12;
Matches 58; Conservative 0; Mismatches 34; Indels 41; Gaps 9;

Query Match 33.8%; Score 235.5; DB 2; Length 432;
 Best Local Similarity 41.7%; Pred. No. 8e-11;
 Matches 58; Conservative 2; Mismatches 32; Indels 47; Gaps 14;

QY 6 HRPPPHGHHP-----HGPPPHGHHP-----HGPPPHGHHP-----HGPPPHGHHP 47
 DB 49 HSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPP 106
 QY 48 PP-----HGHP--HGPP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP 85
 DB 107 PPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPP 166
 QY 86 GHG-----FHDHGCDPPSH 100
 DB 167 KHSPPPPYYHSP-PPPKH 184

RESULT 14
 T10863
 extensin precursor - kidney bean
 N;Alternate names: cell wall protein; hydroxyproline-rich glycoprotein
 C;Species: Phaseolus vulgaris (kidney bean)
 C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 09-Jul-2004
 C;Accession: T10863; B29356
 R;Wycoff, K.L.; Powell, P.A.; Gonzales, R.A.; Corbin, D.R.; Lamb, C.; Dixon, R.A.
 Plant Physiol. 109, 41-52, 1995
 A;Title: Stress activation of a bean hydroxyproline-rich glycoprotein promoter:
 A;Reference number: Z17192; MUID:96061709; PMID:7480331
 A;Accession: T10863
 A;Status: preliminary; translated from GB/EMBL/DDBJ
 A;Molecule type: DNA
 A;Residues: 1-580 <WVC>
 A;Cross-references: UNIPROT:Q09083; EMBL:U18791; NID:g727263; PIDN:AAA87902.1;
 A;Experimental source: strain Tendergreen
 R;Corbin, D.R.; Sauer, N.; Lamb, C.J.
 Mol. Cell. Biol. 7, 4337-4344, 1987
 A;Title: Differential regulation of a hydroxyproline-rich glycoprotein gene fam
 A;Reference number: A29356; MUID:88142825; PMID:3437892
 A;Accession: B29356
 A;Molecule type: mRNA
 A;Residues: 'N',1-229 <COR>
 A;Cross-references: EMBL:M18094; NID:g169346; PIDN:AAA33764.1; PID:g169347
 A;Experimental source: cv. Kievitsboon Koekoek
 C;Superfamily: hydroxyproline-rich glycoprotein
 C;Keywords: glycoprotein; hydroxyproline
 F;1-29/Domain: signal sequence #status predicted <SIG>
 F;30-580/Product: extensin #status predicted <NAR>

Query Match 33.8%; Score 235.5; DB 2; Length 580;
 Best Local Similarity 41.7%; Pred. No. 1e-10;
 Matches 58; Conservative 2; Mismatches 32; Indels 47; Gaps 14;

QY 6 HRPPPHGHHP-----HGPPPHGHHP-----HGPPPHGHHP-----HGPPPHGHHP 47
 DB 61 HSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPP 118
 QY 48 PP-----HGHP--HGPP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP--HGHP 85
 DB 119 PPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPPYYHSHSPPPPKHSGPPPP 178
 QY 86 GHG-----FHDHGCDPPSH 100
 DB 179 KHSPPPPYYHSP-PPPKH 196

RESULT 15
 T51469
 glycine/proline-rich protein - Arabidopsis thaliana
 N;Alternate names: protein K10A8_130
 C;Species: Arabidopsis thaliana (mouse-ear cress)
 C;Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
 C;Accession: T51469

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:32:29 ; Search time 90.8594 Seconds
(without alignments)
569.231 Million cell updates/sec

Title: US-10-074-225A-6

Perfect score: 697

Sequence: 1 SVNIHRPPPHGHPPPP.....HPPHGHGFHDGCPDPSHK 101

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	697	100.0	526	1 HRG RABIT	Q28640 oryctolagus
2	420.5	60.3	546	2 Q98GUL	Q98GUL bos taurus
3	375	53.8	525	2 Q99PS8	Q99PS8 rattus norv
4	333	47.8	525	2 Q99PS5	Q99PS5 mus musculu
5	333	47.8	525	2 Q99PS6	Q99PS6 mus musculu
6	333	47.8	525	2 Q99PSB3	Q99PSB3 mus musculu
7	333	47.8	536	2 Q6YK32	Q6YK32 mus musculu
8	333	47.8	536	2 Q6YKA2	Q6YKA2 mus musculu
9	323	46.3	515	2 Q99PS7	Q99PS7 rattus norv
10	315	45.2	510	2 Q99SB2	Q99SB2 rattus norv
11	314	45.1	312	2 Q68DR3	Q68DR3 homo sapien
12	314	45.1	525	1 Q99HUMAN	Q99HUMAN homo sapien
13	260	37.3	396	1 HRG BOVIN	P33433 bos taurus
14	259.5	37.2	295	2 Q07611	Q07611 rattus norv
15	258.5	37.1	511	2 Q95J00	Q95J00 sus scrofa
16	258.5	37.1	566	2 Q95J01	Q95J01 sus scrofa
17	254.5	35.9	676	2 Q95J09	Q95J09 sus scrofa
18	250	35.9	128	2 Q7M4Q5	Q7M4Q5 homo sapien
19	249	35.7	1109	2 Q7SFA3	Q7SFA3 neurospora
20	247.5	35.5	206	1 PR3 RAT	P04474 rattus norv
21	246.5	35.4	227	2 Q62107	Q62107 mus musculu
22	246.5	35.4	261	1 PR2 MOUSE	P05142 mus musculu
23	246.5	35.4	296	1 PRP3 MOUSE	P05143 mus musculu
24	246.5	35.4	317	2 Q62103	Q62103 mus musculu
25	246	35.3	274	2 Q04154	Q04154 rattus norv
26	240	34.4	1646	2 Q7PRH5	Q7PRH5 anopheles g
27	236	33.9	257	2 Q61517	Q61517 oryza sativ
28	236	33.9	330	2 Q41402	Q41402 sesbania ro
29	235.5	33.8	230	2 Q7DMV8	Q7DMV8 phaseolus v
30	235.5	33.8	392	1 PRP1 HUMAN	P04280 homo sapien
31	235.5	33.8	432	2 Q39835	Q39835 glycine max

32	235.5	33.8	580	2	Q09083	Q09083 phaseolus v
33	234.5	33.6	173	2	Q91F59	Q91F59 arabidopsis
34	233.5	33.5	309	1	NO75 SOYBN	P08297 glycine max
35	233.5	33.5	338	2	Q86FA1	Q86FA1 homo sapien
36	233.5	33.5	424	1	S3B4 HUMAN	Q15427 homo sapien
37	233.5	33.5	424	2	Q8QZY9	Q8QZY9 m splicing
38	233.5	33.5	424	2	Q6AYL5	Q6AYL5 rattus norv
39	233.5	33.5	586	2	Q6CAY8	Q6CAY8 yarrowia li
40	232.5	33.4	1571	2	O54978	O54978 mus musculu
41	232.5	33.4	1572	2	Q8R5N0	Q8R5N0 mus musculu
42	232.5	33.4	1617	2	Q8OU47	Q8OU47 mus musculu
43	231	33.1	173	2	Q15214	Q15214 homo sapien
44	230.5	33.1	382	1	PR82 HUMAN	P02812 homo sapien
45	230	33.0	494	2	Q68D36	Q68D36 homo sapien

ALIGNMENTS

RESULT 1

HRG_RABIT	HRG_RABIT	STANDARD;	PRT;	526 AA.
ID	Q28640;			
AC	Q28640;			
DT	01-NOV-1997 (Rel. 35, Created)			
DT	01-NOV-1997 (Rel. 35, Last sequence update)			
DT	05-JUL-2004 (Rel. 44, Last annotation update)			
DE	Histidine-rich glycoprotein precursor (Histidine-proline rich glycoprotein) (HPRG) (Fragment).			
DE	Name=HRG;			
GN	Oryctolagus cuniculus (Rabbit).			
OS	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.			
OX	NCBI_TaxID=9986;			
RN	[1]			
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.			
RC	TISSUE=Serum;			
RX	MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;			
RA	Borza D.-B., Ratum F.M., Morgan W.T.;			
RT	"Domain structure and conformation of histidine-proline-rich glycoprotein."			
RL	Biochemistry 35:1925-1934(1996).			
CC	!- FUNCTION: The physiological function is not yet known. It binds heme, dyes and divalent metal ions. It can inhibit rosette formation and is known to interact with heparin, thrombospondin, and the lysine-binding site of plasminogen. On the basis of its homology with HMW kinogen, the His-rich region of this protein may mediate the contact activation phase of intrinsic blood coagulation cascade.			
CC	!- SURCELLULAR LOCATION: Secreted.			
CC	!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.			
CC	!- DOMAIN: In addition to having a high His and Pro content, this protein has many internal repeats. 15 tandem repetitions of a 5-residue sequence (G[H/P][H/P]PH, consensus) form a His/Pro-rich region.			
CC	!- SIMILARITY: Contains 2 cystatin-like domains.			

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch.

EMBL; U32189; AAC48516.1; -
InterPro: IPR000010; Prot_inh_cystat.
Pfam: PF00031; Cystatin; I.
Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat;
Signal.
NON TER 1 1
SIGNAL <1 8 Potential.
CHAIN 9 526 Histidine-rich glycoprotein.
DOMAIN 9 126 Cystatin-like 1.

Jones S.J., Marra M.A.; "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences."; Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). [3]

SEQUENCE FROM N.A.

STRAIN=FVB/N; TISSUE=Liver;

Straussberg R.; Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.

EMBL; AB055898; BAB33095.1; -

DR EMBL; BC011168; AAH11168.1; -

DR MGD; MGI:2146636; Hrg.

DR GO; GO:0005615; C:extracellular space; TAS.

DR InterPro; IPK000010; Prot_inh_cstat.

DR Pfam; PF00031; Cystatin; I.

DR SMART; SM00043; CY; 2.

SQ SEQUENCE 525 AA; 59162 MW; A8E93A439CFB126 CRC64;

Query Match 47.8%; Score 333; DB 2; Length 525;
Best Local Similarity 58.3%; Pred. No. 1.2e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIHRPPPHGHG---HP-HGPPPHGHHPGGPPHHGPGPPRRPGRPHGPPRHGPPHHGPPGP 57
Db 322 NRTHR-PSYNSMCNEPCNHRPHGHHPSHHPFGHHSHG----HHPHGHHPHSHHSKH 376

Qy 58 PPHGHPHGPPPHGPPHHGPPHHGPHGHGFHDHGCDPPSH 100
Db 377 HPPGHHPHGHHPHHGHHPHGHHPHHGHDFLDYGCDPPSN 419

RESULT 5

Q99PS6 Q99PS6 PRELIMINARY; PRT; 525 AA.

AC Q99PS6;

DT 01-JUN-2001 (TrEMBLrel. 17, Created)

DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)

DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)

DE Histidine-rich glycoprotein.

DR Name=Hrg; Synonyms=MMHRG;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10090;

[1]

SEQUENCE FROM N.A.

STRAIN=BALB/c; TISSUE=Liver;

Ra Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
Ra Onishi T., Sato N., Hirata D., Teuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DDBJ databases.

DR EMBL; AB055897; BAB33094.1; -

DR MGD; MGI:2146636; Hrg.

DR GO; GO:0005615; C:extracellular space; TAS.

DR InterPro; IPK000010; Prot_inh_cstat.

DR Pfam; PF00031; Cystatin; I.

DR SMART; SM00043; CY; 2.

SQ SEQUENCE 525 AA; 59090 MW; A8E93A439CFB3AC CRC64;

Query Match 47.8%; Score 333; DB 2; Length 525;
Best Local Similarity 58.3%; Pred. No. 1.2e-14;
Matches 60; Conservative 3; Mismatches 30; Indels 10; Gaps 4;

Qy 3 NIHRPPPHGHG---HP-HGPPPHGHHPGGPPHHGPGPPRRPGRPHGPPRHGPPHHGPPGP 57
Db 322 NRTHR-PSYNSMCNEPCNHRPHGHHPSHHPFGHHSHG----HHPHGHHPHSHHSKH 376

Qy 58 PPHGHPHGPPPHGPPHHGPPHHGPHGHGFHDHGCDPPSH 100
Db 377 HPPGHHPHGHHPHHGHHPHGHHPHHGHDFLDYGCDPPSN 419

RESULT 6

Q9ESB3

[illegible]

```

Query Match      45.1%; Score 314; DB 2; Length 312;
Best Local Similarity 41.8%; Pred. No. 1.4e-13;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

QY 7 RPP--PHG---HHPHGPPPHGHHPHGPPPP-----HGHP-PHGPPP----- 40
DB 62 KPFPKPHGSRDHHHPKP-----HEHGPPPPPPDERDHSHPPLPQGPPPLPMSCSQOH 116
QY 41 -----RHP-----PHGPPPHGHP-----HGPPPHGHPHPGP 67
DB 117 ATFGTNGAQRSHNNSSDLPHPKHSHSQHPGHGHPHAAHPHEHDTHRQHPGHGHPGH 176
QY 68 PPHGHPHPGPPPHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGH 100
DB 177 HPHGHPHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHP 209

RESULT 12
HRG_HUMAN
ID HRG_HUMAN STANDARD; PRT; 525 AA.
AC P04196;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein precursor (Histidine-proline rich
DE glycoprotein) (HPRG).
GN Names=HRG;
OS Homo sapiens (Human).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86216149; PubMed=3011081;
RA Koide T., Foster D.C., Yoshitake S., Davie E.W.;
RT "Amino acid sequence of human histidine-rich glycoprotein derived from
RT the nucleotide sequence of its cDNA.";
RL Biochemistry 25:2220-2225(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 214-247 FROM N.A.
RX MEDLINE=94245171; PubMed=8188234;
RA Hennis B.C., Franks R.R., Bakker E., Vossen R.H., van der Poort E.W.,
RA Blonden L.A., Cox S., Khan P.M., Spurr N.K., Kluff C.;
RT "Evidence for the absence of intron H of the histidine-rich
RT glycoprotein (HRG) gene: genetic mapping and in situ localization of
RT HRG to chromosome 3q28-q29.";
RL Genomics 19:195-197(1994).
RN [4]
RP SEQUENCE OF 19-27.
RC TISSUE=Plasma;
RX MEDLINE=93092937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.P.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992).
CC -1- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the blood coagulation phase of intrinsic blood
CC coagulation cascade.
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -1- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 12 tandem repetitions of a 5-

```

```

CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -1- SIMILARITY: Contains 2 cystatin-like domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
CC EMBL: M13149; AAA52694.1; --
CC EMBL: AB005803; BAA21613.1; --
CC EMBL: Z17218; CAA78925.1; --
CC FIR; A01287; KGHUGH.
CC SWISS-2DPAGE; P04196; HUMAN.
CC Genew; HGNC:5181; HRG.
CC MIM; 142640; --
CC InterPro; IPR000010; Prot_inh_cystat.
CC Pfam; PF00311; Cystatin; I.
KW Direct protein sequencing; Glycoprotein; Heparin-binding;
KW Polymorphism; Repeat; Signal.
FT SIGNAL 1 18 Histidine-rich glycoprotein.
FT CHAIN 19 525 Cystatin-like 1.
FT DOMAIN 137 254 Cystatin-like 2.
FT DOMAIN 276 321 Pro-rich.
FT DOMAIN 350 497 His/Pro-rich.
FT DISULFID 24 504 By similarity.
FT DISULFID 78 89 By similarity.
FT DISULFID 105 126 By similarity.
FT DISULFID 203 417 By similarity.
FT DISULFID 218 241 By similarity.
FT CARBOHYD 63 63 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 125 125 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 344 344 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 345 345 N-linked (GlcNAc...) (Potential).
FT VARIANT 204 204 P -> S (in dbSNP:3181917).
FT FTID=VAR 014528.
SQ SEQUENCE 525 AA; 59578 MW; A2B124D6CE93114F CRC64;

Query Match      45.1%; Score 314; DB 1; Length 525;
Best Local Similarity 41.8%; Pred. No. 2.1e-13;
Matches 64; Conservative 2; Mismatches 23; Indels 64; Gaps 8;

QY 7 RPP--PHG---HHPHGPPPHGHHPHGPPPP-----HGHP-PHGPPP----- 40
DB 275 KPFPKPHGSRDHHHPKP-----HEHGPPPPPPDERDHSHPPLPQGPPPLPMSCSQOH 329
QY 41 -----RHP-----PHGPPPHGHP-----HGPPPHGHPHPGP 67
DB 330 ATFGTNGAQRSHNNSSDLPHPKHSHSQHPGHGHPHAAHPHEHDTHRQHPGHGHPGH 389
QY 68 PPHGHPHPGPPPHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGH 100
DB 390 HPHGHPHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHPGHGHP 422

RESULT 13
HRG_BOVIN
ID HRG_BOVIN STANDARD; PRT; 396 AA.
AC P33433;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein (Histidine-proline rich glycoprotein)
DE (HPRG) (Fragments).
GN Name=HRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.

```


Job time : 92.8594 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-7

Perfect score: 23

Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesepc_16Dec04:*

- 1: Genesepc1980a:*
- 2: Genesepc1990a:*
- 3: Genesepc2000a:*
- 4: Genesepc2001a:*
- 5: Genesepc2002a:*
- 6: Genesepc2003a:*
- 7: Genesepc2003ba:*
- 8: Genesepc2004a:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	91.3	4	5	Abg77500 Targettin
2	21	91.3	5	2	Aar12701 Pentapept
3	21	91.3	5	2	Aar51565 Mimotope
4	21	91.3	5	2	Aar69933 Pentameri
5	21	91.3	5	2	Aar89833 Melanotro
6	21	91.3	5	2	Aar98679 Peptide 5
7	21	91.3	5	2	Aaw78934 Antidepre
8	21	91.3	5	2	Aay23419 V beta 6
9	21	91.3	5	5	Abb79810 Histidine
10	21	91.3	5	5	Abb79808 Histidine
11	21	91.3	5	5	Abb79809 Histidine
12	21	91.3	5	5	Abb79811 Histidine
13	21	91.3	5	5	Aag80338 Enzyme cl
14	21	91.3	5	5	Aao15016 Mutant VL
15	21	91.3	5	8	Adh10419 HPRG prot
16	21	91.3	5	8	Adh10422 HPRG prot
17	21	91.3	5	8	Adh10417 Rabbit HP
18	21	91.3	5	8	Adh10415 Rabbit HP
19	21	91.3	5	8	Adh10416 Rabbit HP
20	21	91.3	5	8	Adh10418 HPRG prot
21	21	91.3	5	8	Ado81122 Sheep pri
22	21	91.3	5	8	Adr03627 E. coli m
23	21	91.3	6	2	Aaw75370 Hexapepti
24	21	91.3	6	2	Aaw75302 Hexapepti
25	21	91.3	6	4	Aab55516 Human ela

26	21	91.3	6	5	ABG31535	Abg31535 Camptothe
27	21	91.3	6	5	AAE26391	Aae26391 Human GPR
28	21	91.3	6	5	ABP53526	Abp53526 Camptothe
29	21	91.3	6	7	ADN12399	Adn12399 Peptide #
30	21	91.3	6	8	ADO37172	Ado37172 Binding p
31	21	91.3	6	8	ADO37904	Ado37904 Binding p
32	21	91.3	6	8	ADO37705	Ado37705 Binding p
33	21	91.3	6	8	ADO37180	Ado37180 Binding p
34	21	91.3	6	8	ADO37698	Ado37698 Binding p
35	21	91.3	6	8	ADO37897	Ado37897 Binding p
36	21	91.3	6	8	ADO28227	Ado28227 Capture s
37	21	91.3	6	8	ADO28419	Ado28419 Capture s
38	21	91.3	6	8	ADO28426	Ado28426 Capture s
39	21	91.3	6	8	ADO27694	Ado27694 Capture s
40	21	91.3	6	8	ADO27702	Ado27702 Capture s
41	21	91.3	6	8	ADO28220	Ado28220 Capture s
42	21	91.3	6	8	ADR49685	Adr49685 6-mer pep
43	21	91.3	6	8	ADR50404	Adr50404 6-mer pep
44	21	91.3	6	8	ADR49879	Adr49879 6-mer pep
45	21	91.3	6	8	ADR49871	Adr49871 6-mer pep

ALIGNMENTS

RESULT 1

ABG77500
ID ABG77500 standard; peptide; 4 AA.
XX

AC ABG77500;
XX

DT 05-NOV-2002 (first entry)
XX

DE Targetting peptide selective for human organ, tissue or cell type #33.
XX

KW Human; cytostatic; antiinflammatory; antidiabetic; cardiovascular;
KW immunomodulator; antibacterial; antiviral; gene therapy; cancer;
KW arthritis; diabetes; inflammatory disease; atherosclerosis;
KW autoimmune disease; bacterial infection; viral infection;
KW cardiovascular disease; degenerative disease.
XX

OS Homo sapiens.
XX

PN WO200220723-A2.
XX

PD 14-MAR-2002.
XX

PF 07-SEP-2001; 2001WO-US028044.
XX

PR 08-SEP-2000; 2000US-0231266P.
XX

PR 17-JAN-2001; 2001US-00765101.
XX

PA (TEXA) UNIV TEXAS SYSTEM.
XX

PI Arap W, Pasqualini R;
XX

PT WPI; 2002-599247/64.
XX

DR New targeting peptides identified by phage display, useful for treating a disease state, e.g. cancer, diabetes, inflammatory disease, atherosclerosis, autoimmune disease, bacterial infection, viral infection, cardiovascular disease or degenerative disease. This sequence represents a human targeting peptide selective for human organs, tissues or cell types
XX

PS Claim 16; Page 63; 269pp; English.
XX

CC The invention describes an isolated peptide of 100 amino acids or less in size. The peptide is useful for treating a disease state, e.g. cancer, arthritis, diabetes, inflammatory disease, atherosclerosis, autoimmune disease, bacterial infection, viral infection, cardiovascular disease or degenerative disease. This sequence represents a human targeting peptide selective for human organs, tissues or cell types
XX

Sequence 4 AA;

Query Match 91.3%; Score 21; DB 5; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 1 PHG 3

RESULT 2
AAR12701
ID AAR12701 standard; protein; 5 AA.

XX AAR12701;
AC AAR12701;
XX 31-JUL-1991 (first entry)

XX Pentapeptide paralogue for use as chromatographic affinity ligand.
XX HPLC.
XX Synthetic.

XX WO9106356-A.
XX 16-MAY-1991.
XX 31-OCT-1989; 89US-00429721.

XX 31-OCT-1989; 89US-00429721.
XX (TERR-) TERRAPIN TECHN INC.
XX Kauvar LM;

XX WPI; 1991-163985/22.
XX Identifying paralogues with specific affinity for analyte - using
XX candidate paralogues with systematically varied values of at least 2
XX parameters.

XX Claim 41; Fig 8; 95pp; English.
XX Peptide is one of a panel, each of which have systemically varied values of
XX at least two parameters. A test protein is matched against the panel, and
XX the paralogues with highest binding affinity are selected. Selected
XX paralogues are bound to a matrix which is incorporated into the HPLC gel
XX and used as affinity ligands. Technique is useful in chromatographic
XX separation, purification and binding assay

QY 3 PHG 5
Db 2 PHG 4

Query Match 91.3%; Score 21; DB 2; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 2 PHG 4

RESULT 3
AAR51565
ID AAR51565 standard; peptide; 5 AA.

XX AAR51565;
XX 25-MAR-2003 (revised)
XX 19-OCT-1994 (first entry)

XX Mimotope peptide #59 from panel of maximally diverse mimotopes.
XX

KW mimotope panel; rational drug design; candidate drug; screening assay;
KW hydrophobicity; antibody repertoire.
XX Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1
/note= "acetylated"

XX US5300425-A.
XX 05-APR-1994.

XX 06-DEC-1989; 89US-00447009.
XX 13-OCT-1987; 87US-00108130.
XX 11-OCT-1988; 88US-00255906.

XX (TERR-) TERRAPIN TECHNOLOGIES INC.
XX Kauvar LM;

XX WPI; 1994-109390/13.
XX Screening of candidate drugs for binding to receptor - by comparing
XX inverse image antibody profile of drug with mimotope-binding profile of
XX receptor.

XX Example 3; Fig 3; 29pp; English.
XX A panel of 88 pentapeptides was designed on the basis of decreasing
XX hydrophobicity and periodic variation of hydrophobic moment. The peptides
XX were labelled with iodine-125 and tested with individual members of a
XX basal antibody repertoire. (Spleen cells were harvested from mice and
XX used to provide a panel of antibody-secreting hybridoma cells as a subset
XX of the complete B-cell repertoire). Nearly uniform binding to all
XX antibody members of the repertoire was observed. The test was then
XX repeated with the addition of a defined amount of analyte to the mixture.
XX A small number of wells showed greatly decreased labelling and these
XX antibodies represented the successful result of an initial screen for
XX those which preferentially bind analyte. (Updated on 25-MAR-2003 to
XX correct PF field.)

XX Sequence 5 AA;
XX Query Match Similarity 91.3%; Score 21; DB 2; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 2 PHG 4

RESULT 4
AAR69933
ID AAR69933 standard; peptide; 5 AA.

XX AAR69933;
XX 25-MAR-2003 (revised)
XX 18-OCT-1995 (first entry)

XX Pentameric mimotope 59 used to obtain highly specific antibodies.
XX mimotope; antibody; production; high specificity; detection; immunoassay;
XX high performance liquid chromatography.
XX Synthetic.

XX Key Location/Qualifiers
FT Modified-site 1
/note= "Acetyl-His"

PN US5384263-A.
 XX 24-JAN-1995.
 PD
 XX
 XX 04-JUN-1993; 93US-00072190.
 XX
 XX 13-OCT-1987; 87US-00108130.
 PR 11-OCT-1988; 88US-00255906.
 XX
 XX (TERR-) TERRAPIN TECHNOLOGIES INC.
 PA
 XX Kauvar LM;
 XX WPI; 1995-105497/14.
 DR
 XX
 XX Producing antibodies with high specificity and affinity for an analyte -
 PT by immunisation with selected mimotope, also analyte detection kits,
 PT useful for immunoassay of materials usually analysed by HPLC.
 XX
 XX Example 3; Fig 3; 25pp; English.
 PS
 XX AAR69875-969 are pentameric mimotopes designed on the basis of decreasing
 CC hydrophobicity and periodic variation of hydrophobic moment. All the
 CC pentapeptides are acetylated at the N-terminus except E12 (for FITC)
 CC (sic). E12 is not identified in the specification. The mimotopes are used
 CC in the method of the invention to obtain antibodies specifically and
 CC strongly reactive with a desired analyte. The mimotope is obtd. by
 CC reacting a panel of starting antibodies (Abs) representative of the
 CC reacting B cell repertoire of a mammal with an analyte (so as to identify
 CC analyte-reacting Abs) and then reacting each of a panel of candidate
 CC mimotopes representative of a random set of 3D contours with the analyte-
 CC reacting Abs. A subject is immunised with one or more mimotopes
 CC identified and the product Abs are recovered from the serum of the
 CC subject. (Updated on 25-MAR-2003 to correct PF field.)
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 91.3%; Score 21; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 DB ||||
 ||||
 2 PHG 4
 RESULT 5
 AAR89833
 ID AAR89833 standard; peptide; 5 AA.
 XX
 XX AAR89833;
 AC
 XX 18-JUN-1996 (first entry)
 DT
 XX
 DE Melanotrophic release inhibiting factor (MIF) modified peptide analogue.
 XX
 XX MIF; melanotrophic release inhibiting factor; modified; analogue;
 KW hormone; depression; low dosage; L-propyl L-leucyl glycine;
 KW melanocyte stimulating inhibitory factor.
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /label= OTHER
 FT /note= "4-F-Phe"
 FT
 FT Modified-site 2 /label= 4Hyp
 FT /note= "4-OH-Pro, may be in cis- or trans- form"
 FT Modified-site 5 /label= OTHER
 FT /note= "amidated"
 FT

XX WO9530430-A1.
 PN
 XX 16-NOV-1995.
 PD
 XX
 XX 02-MAY-1995; 95WO-US005560.
 XX
 XX 04-MAY-1994; 94US-00238089.
 PR
 XX (INNA-) INNAPHARMA INC.
 PA
 XX Abajian HB, Noble JF, Hlavka JJ;
 PI WPI; 1995-403936/51.
 XX
 XX New tri-, tetra-, penta- and poly-peptide cpds. - based on the
 PT tri-peptide hormone melanocyte stimulating inhibitory factor, used for
 PT treating depression.
 PT
 XX Claim 63; Page 32; 138pp; English.
 PS
 XX AAR9786-R89858 are small peptide compounds based on the tripeptide
 CC hormone melanotrophic release inhibiting factor (MIF), also known as L-
 CC propyl L-leucyl glycine and melanocyte stimulating inhibitory factor. The
 CC peptides are used for treating depression and can be administered so as
 CC to achieve a circulating plasma level of 30-90 mg per average human
 CC adult. The peptides show greater anti-depressant activity than MIF. They
 CC can be administered at lower dosage than known anti-depressants which
 CC reduces potential deleterious side effects
 XX
 XX Sequence 5 AA;
 SQ
 Query Match 91.3%; Score 21; DB 2; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 DB ||||
 ||||
 2 PHG 4
 RESULT 6
 AAR98679
 ID AAR98679 standard; peptide; 5 AA.
 XX
 XX AAR98679;
 AC
 XX 25-MAR-2003 (revised)
 DT 11-MAR-1997 (first entry)
 DT
 XX Peptide 59 from 88 member diverse mimotope panel.
 DE
 XX Panel; mimotope; decreasing hydrophobicity; periodic variation;
 KW hydrophobic moment; antibody; repertoire; identification; drug;
 KW candidate; receptor; binding; ligand; rational; design; selection;
 KW treatment; tumour; production; immunological reagent; analyte; detection;
 KW trace contaminant; mimotope.
 XX
 XX Synthetic.
 OS
 XX
 XX Key Location/Qualifiers
 FH Modified-site 1 /note= "acylated"
 FT
 FT US5541070-A.
 PN
 XX 30-JUL-1996.
 PD
 XX
 XX 08-SEP-1993; 93US-00118133.
 PF
 XX 13-OCT-1987; 87US-00108130.
 PR 11-OCT-1988; 88US-00255906.
 PR 06-DEC-1989; 89US-00447009.
 PR

```

XX (KAUV/) KAUVAR L M.
XX KAUVAR LM;
XX WPI; 1996-361955/36.
XX
XX Identifying candidate drugs that bind a specific receptor by competitive
XX reaction with panel of mimotope(s) - useful in rational drug design.
XX
XX Example 3; Fig 3; 27pp; English.
XX
XX The present peptide is a member of a panel of 88 pentapeptide mimotopes
XX designed on the basis of decreasing hydrophobicity, and periodic
XX variation of hydrophobic moment. The panel was synthesised using the
XX method of Geyesen, H. M., et al, Proc. Natl. Acad. Sci. USA (1984), which
XX uses lots of 96 pins; the remaining 8 polylethylene pins being controls.
XX The mimotopes were then mixed, 125-I labelled and tested with individual
XX members of a basal antibody (Ab) repertoire. Nearly uniform binding to
XX all members was found. The test was then repeated with the addition of a
XX defined amt. of analyte. A small number showed greatly increased
XX labelling, these Ab representing the successful result of an initial
XX screen for those that pref. bind analyte. The above is an example of a
XX claimed method for identifying members of a panel of candidate drugs,
XX that bind to a receptor having a known ligand. It is useful in rational
XX drug design, e.g. selection of monoclonal Ab for treating individual
XX tumours, and for the prodn. of immunological reagents for any analyte,
XX including those not normally detectable by immunoassay, e.g. trace
XX contaminants in soil, air or water. (Updated on 25-MAR-2003 to correct PF
XX field.)
XX
XX Sequence 5 AA;
XX
XX Query Match 91.3%; Score 21; DB 2; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3 PHG 5
Db |||
2 PHG 4

RESULT 7
AAW78934
ID AAW78934 standard; peptide; 5 AA.
AC AAW78934;
XX
XX 20-NOV-1998 (first entry)
XX
XX Antidepressant MIF based peptide #64.
XX
XX Antidepressant; depression; tripeptide MIF; hormone; bovine; rat;
XX melanocyte stimulating inhibitory factor; psychiatric disorder;
XX neurochemical; central nervous system.
XX
XX Synthetic.
XX Bos taurus.
XX Rattus sp.
XX
XX Key Location/Qualifiers
XX Modified-site 1 /note= "4F-Phe"
XX Modified-site 2 /label= 4HYP
XX Modified-site 5 /note= "4-hydroxyproline"
XX
XX US5767083-A.
XX
XX 16-JUN-1998.
XX
XX
XX Query Match 91.3%; Score 21; DB 2; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3 PHG 5
Db |||
2 PHG 4

RESULT 8
AAW23419
ID AAW23419 standard; peptide; 5 AA.
AC AAW23419;
XX
XX 02-SEP-1999 (first entry)
XX
XX V beta 6 clone found in MS patients after vaccination with TCR.
XX
XX Vaccine; T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;
XX V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;
XX multiple sclerosis.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9927957-A1.
XX
XX 10-JUN-1999.
XX
XX 03-DEC-1997; 97WO-US023147.
XX
XX 03-DEC-1997; 97WO-US023147.
XX
XX (IMMU-) IMMUNE RESPONSE CORP.
XX (KIMM-) KIMMEL CANCER CENT SIDNEY.
XX
XX Brostoff SW, Wilson DB, Smith LR, Gold DP, Carlo DJ;
XX WPI; 1999-404801/34.
XX
XX T0 cell receptor peptide-derived vaccines.
XX
XX Example 11; Page 85; 104pp; English.
XX
XX
XX PF 02-MAY-1995; 95US-00432651.
XX PR 04-MAY-1994; 94US-00238089.
XX PA (INNA-) INNAPHARMA INC.
XX PI Hlavka JJ, Noble JF, Abajian HB;
XX DR WPI; 1998-361753/31.
XX
XX New peptides based on structure of MIF - useful for treatment of
XX depression.
XX
XX Claim 14; Col 99; 56pp; English.
XX
XX Peptides given in AAW78922 to AAW79000 are based on the tripeptide MIF
XX (melanocyte stimulating inhibitory factor), which was initially isolated
XX from bovine and rat hypothalamic extracts. The peptides from the present
XX invention are useful for treating depression, optionally together with
XX other antidepressants e.g. fluoxetine, amitriptyline or sertraline. The
XX peptides avoid the serious side effects of the prior art tricyclics, MAO
XX oxidase inhibitors and second generation antidepressants. They are also
XX lipid soluble and easily cross the blood-brain barrier, but are not so
XX easily metabolised as the parent MIF (the tripeptide Pro-Leu-Gly-NH2) and
XX so allows a lower dosage
XX
XX Sequence 5 AA;
XX
XX Query Match 91.3%; Score 21; DB 2; Length 5;
XX Best Local Similarity 100.0%; Pred. No. 1.8e+06;
XX Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 3 PHG 5
Db |||
2 PHG 4

RESULT 8
AAW23419
ID AAW23419 standard; peptide; 5 AA.
AC AAW23419;
XX
XX 02-SEP-1999 (first entry)
XX
XX V beta 6 clone found in MS patients after vaccination with TCR.
XX
XX Vaccine; T cell receptor; TCR; T cell; V beta 6.2/3; V beta 6/5;
XX V beta 6.7; V beta 2; V beta 5/1; V beta 7; V beta 13; V beta 6;
XX multiple sclerosis.
XX
XX Synthetic.
XX Homo sapiens.
XX
XX WO9927957-A1.
XX
XX 10-JUN-1999.
XX
XX 03-DEC-1997; 97WO-US023147.
XX
XX 03-DEC-1997; 97WO-US023147.
XX
XX (IMMU-) IMMUNE RESPONSE CORP.
XX (KIMM-) KIMMEL CANCER CENT SIDNEY.
XX
XX Brostoff SW, Wilson DB, Smith LR, Gold DP, Carlo DJ;
XX WPI; 1999-404801/34.
XX
XX T0 cell receptor peptide-derived vaccines.
XX
XX Example 11; Page 85; 104pp; English.
XX
XX

```

CC The specification describes vaccines which comprise immunologically
 CC effective amounts of T cell receptor (TCR) peptides. The TCRs are present
 CC on the surface of T cells. The TCRs are chosen from V beta 6.2/3, V beta
 CC 6/5, V beta 6.7, V beta 2, V beta 5/1, V beta 7 or V beta 13. The V beta
 CC TCR peptide-based vaccines are useful for prevention or treatment of
 CC multiple sclerosis (MS). The presence of V beta 6.7 appears to be
 CC particularly associated with multiple sclerosis and can be used to
 CC determine an individual's susceptibility to multiple sclerosis.
 CC Vaccinating, rather than passively administering heterologous antibodies,
 CC allows the host's own immune system to mobilize and suppress auto
 CC aggressive T cells. Therefore, the suppression is persistent and may
 CC involve any and all immunological mechanisms in effecting that
 CC suppression. Such a multi-faceted response is more effective than the uni
 CC dimensional suppression achieved by passive administration of monoclonal
 CC antibodies or extant-derived regulatory T cell clones. AA23387-Y23480
 CC represent peptides derived from TCR V beta 6 clones found in the
 CC cerebrospinal fluid (CSF) of MS patients, after vaccination with V beta 6
 CC
 CC Sequence 5 AA;
 CC
 CC Query Match 91.3%; Score 21; DB 2; Length 5;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 CC Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 3 PHG 5
 CC Db |||
 CC 1 PHG 3
 CC
 CC RESULT 9
 CC ABB79810
 CC ID ABB79810 standard; peptide; 5 AA.
 CC AC ABB79810;
 CC DT 25-NOV-2002 (first entry)
 CC DE Histidine proline rich glycoprotein pentapeptide.
 CC
 CC KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
 CC antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
 CC antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
 CC neurotropic; neuroprotective; antiparkinsonian.
 CC
 CC OS Synthetic.
 CC
 CC PN WO200264621-A2.
 CC
 CC PD 22-AUG-2002.
 CC
 CC PF 14-FEB-2002; 2002WO-US004336.
 CC
 CC PR 14-FEB-2001; 2001US-0268370P.
 CC
 CC PA (ATTE-) ATTENUON LLC.
 CC
 CC PI Donate F, Harris S, Plunkett ML, Mazar AP;
 CC
 CC XX WPI; 2002-666989/71.
 CC DR
 CC PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
 CC diagnosing or treating diseases associated with undesired cell migration,
 CC invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
 CC
 CC PS Claim 2; Page 67; 82pp; English.
 CC
 CC XX The present sequence is a specific example of claimed anti-angiogenic
 CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
 CC peptides comprise: the histidine-proline-rich (H/P) domain of human
 CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
 CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
 CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
 CC tube formation in vitro or in vivo; or a pentapeptide having the generic

CC sequence given in ABB79808, such as the present peptide, or its variant
 CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
 CC N- or C-terminus. Also claimed are: chemically synthesised or
 CC recombinantly produced peptide multimers; a diagnostically or
 CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
 CC multimer; a diagnostically useful HPRG-related composition, comprising
 CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
 CC carrier; an antibody specific for an epitope of HPRG that is present in
 CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
 CC of the domains in a way which inhibits the anti-angiogenic activity of
 CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
 CC method for inhibiting cell migration, cell invasion, cell proliferation
 CC or angiogenesis, or for inducing apoptosis; a method for treating a
 CC subject having a disease or condition associated with undesired cell
 CC migration, invasion, proliferation, or angiogenesis; a method of
 CC stimulating or inhibiting angiogenesis in a subject; a method of
 CC detecting the presence of HPRG or its cleavage product or its peptide in
 CC a biological sample; isolated nucleic acids encoding the polypeptide,
 CC peptide or peptide multimer; an expression vector; transformed or
 CC transfected cells; a method of providing to a cell, tissue or organ an
 CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
 CC pentapeptide, or the peptide multimer; an affinity ligand useful for
 CC binding to, or isolating, an HPRG-binding molecule or cells expressing
 CC the binding molecule, comprising the polypeptide, peptide or peptide
 CC multimer, immobilised to a solid support or carrier; and a method of
 CC isolating HPRG-binding molecule, or isolating or enriching cells
 CC expressing HPRG-binding site or receptor, from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays
 CC
 CC XX Sequence 5 AA;
 CC
 CC Query Match 91.3%; Score 21; DB 5; Length 5;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 CC Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 3 PHG 5
 CC Db |||
 CC 3 PHG 5
 CC
 CC RESULT 10
 CC ABB79808
 CC ID ABB79808 standard; peptide; 5 AA.
 CC AC ABB79808;
 CC XX
 CC DT 25-NOV-2002 (first entry)
 CC DE Histidine proline rich glycoprotein pentapeptide.
 CC
 CC KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
 CC antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
 CC antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
 CC neurotropic; neuroprotective; antiparkinsonian.
 CC
 CC OS Synthetic.
 CC
 CC XX Key Location/Qualifiers
 CC FT Misc-difference 1 /label= His, Pro
 CC FT Misc-difference 2 /label= His, Pro
 CC
 CC XX WO200264621-A2.

CC expressing HPRG-binding site or receptor, from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays
 CC
 CC Sequence 5 AA;
 CC
 CC Query Match 91.3%; Score 21; DB 5; Length 5;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 CC Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 3 PHG 5
 CC |||
 CC Db 3 PHG 5
 CC
 CC RESULT 12
 CC ID ABB79811
 CC XX ABB79811 standard; peptide; 5 AA.
 CC AC ABB79811;
 CC XX
 CC DT 25-NOV-2002 (first entry)
 CC XX
 CC DE Histidine proline rich glycoprotein pentapeptide.
 CC XX
 CC KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
 CC KW antiatherosclerotic; antiinflammatory; antidiabetic; gynaecological;
 CC KW antarthritic; anticulcer; osteopathic; antitumour; ophthalmological;
 CC KW neurotropic; neuroprotective; antiparkinsonian.
 CC XX
 CC OS Synthetic.
 CC
 CC PN WO200264621-A2.
 CC XX
 CC PD 22-AUG-2002.
 CC XX
 CC PF 14-FEB-2002; 2002WO-US004336.
 CC XX
 CC PR 14-FEB-2001; 2001US-0268370P.
 CC XX
 CC PA (ATTE-) ATTENUON LLC.
 CC XX
 CC PI Donate F, Harris S, Plunkett ML, Mazar AP;
 CC XX
 CC DR WPI; 2002-666989/71.
 CC XX
 CC PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
 CC diagnosing or treating diseases associated with undesired cell migration,
 CC PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
 CC XX
 CC PS Claim 2; Page 67; 82pp; English.
 CC
 CC The present sequence is a specific example of claimed anti-angiogenic
 CC pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
 CC peptides comprise: the histidine-proline-rich (H/P) domain of human
 CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
 CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
 CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
 CC tube formation in vitro or in vivo; or a pentapeptide having the generic
 CC sequence given in ABB79808, such as the present peptide, or its variant
 CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
 CC N- or C-terminus. Also claimed are: chemically synthesised or
 CC recombinantly produced peptide multimers; a diagnostically or
 CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
 CC multimer; a diagnostically useful HPRG-related composition, comprising

CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
 CC carrier; an antibody specific for an epitope of HPRG that is present in
 CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
 CC of the domains in a way which inhibits the anti-angiogenic activity of
 CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
 CC method for inhibiting cell migration, cell invasion, cell proliferation
 CC or angiogenesis, or for inducing apoptosis; a method for treating a
 CC subject having a disease or condition associated with undesired cell
 CC migration, invasion, proliferation, or angiogenesis; a method of
 CC stimulating or inhibiting angiogenesis in a subject; a method of
 CC detecting the presence of HPRG or its cleavage product or its peptide in
 CC a biological sample; isolated nucleic acids encoding the polypeptide,
 CC peptide or peptide multimer; an expression vector; transformed or
 CC transfected cells; a method of providing to a cell, tissue or organ an
 CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
 CC pentapeptide, or the peptide multimer; an affinity ligand useful for
 CC binding to, or isolating, an HPRG-binding molecule or cells expressing
 CC the binding molecule, comprising the polypeptide, peptide or peptide
 CC multimer, immobilised to a solid support or carrier; and a method of
 CC isolating HPRG-binding molecule, or isolating or enriching cells
 CC expressing HPRG-binding site or receptor, from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays
 CC
 CC Sequence 5 AA;
 CC
 CC Query Match 91.3%; Score 21; DB 5; Length 5;
 CC Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 CC Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC
 CC QY 3 PHG 5
 CC |||
 CC Db 3 PHG 5
 CC
 CC RESULT 13
 CC ID AAG80338
 CC XX AAG80338 standard; peptide; 5 AA.
 CC AC AAG80338;
 CC XX
 CC DT 20-FEB-2002 (first entry)
 CC XX
 CC DE Enzyme cleavable peptide SEQ ID 18.
 CC XX
 CC KW Antineoplastic; drug; cytostatic; chemotherapeutic agent; matrixin;
 CC KW cancer; leukaemia; lymphoma; carcinoma; sarcoma; melanoma.
 CC XX
 CC OS Synthetic.
 CC
 CC Key Location/Qualifiers
 CC FT Modified-site 1 /note= "Acetylated residue"
 CC FT Modified-site 4 /label= OTHER
 CC FT Modified-site 5 /note= "Homophenylalanine, Hof"
 CC FT Modified-site 5 /label= OTHER
 CC FT /note= "Residue is conjugated to Dox, where Dox is the
 CC anthracycline doxorubicin"
 CC
 CC WO200168145-A2.
 CC
 CC 20-SEP-2001.
 CC
 CC XX

PF	15-MAR-2001; 2001WO-US008589.	PT	ribosome display system.
XX		XX	
PR	15-MAR-2000; 2000US-0189387P.	PS	Example 9; Page 49; 61pp; English.
XX		XX	
PA	(DUPO) DUPONT PHARM CO.	CC	The invention comprises a method (ribosome display system) for obtaining
XX		CC	a specific binding pair (sbp) member that binds a complementary sbp
PI	Copeland RA, Albright CF, Combs AP, Dowling RL, Graciani NR;	CC	member of interest. The method involves incubating mRNA molecules
PI	Han W, Higley CA, Huang PS, Yue EW, Dimeo SV;	CC	encoding an sbp and lacking an in-frame stop codon, allowing ribosome
XX		CC	translation of the mRNA to produce the encoded sbp member, forming
DR	WPI; 2002-010546/01.	CC	complexes comprising ribosome, mRNA, and encoded sbp member displayed on
XX		CC	the ribosome. The ribosome display system is useful for the selection of
PT	New compounds comprising enzyme cleavable peptide conjugated to	CC	an sbp member able to bind a complementary sbp member. The present amino
PT	antineoplastic agent used as chemotherapeutic agents in targeted	CC	acid sequence represents a mutant VL CDR3 peptide
PT	treatment of cancers.	XX	
XX		XX	Sequence 5 AA;
PS	Claim 33; Page 112; 203pp; English.		
XX		Query Match	91.3%; Score 21; DB 5; Length 5;
CC	This invention relates to novel compounds (I) comprising antineoplastic	Best Local Similarity	100.0%; Pred. No. 1.8e+06;
CC	agents conjugated to enzyme cleavable peptides comprising the amino acid	Matches	3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
CC	recognition sequence of a membrane bound and/or cell-secreted peptidase.		
CC	The compounds of the invention have cytostatic activity. The products of	QY	3 PHG 5
CC	the invention can be used as chemotherapeutic agents, preferably in the	DB	2 PHG 4
CC	presence of a matrixin, used for targeted treatment of cancers,		
CC	particularly leukaemia, lymphoma, carcinoma, sarcoma and melanoma. (I)		
CC	are stable to non-specific enzymatic degradation in the blood stream and		
CC	reduce damage to healthy, non-target tissue, so that the toxicity of (I)		
CC	is reduced, compared to its non-conjugated form. This sequence represents	RESULT 15	
CC	an enzyme cleavable peptide described in the method of the invention	ADH10419	
XX		ID	ADH10419 standard; peptide; 5 AA.
XX		AC	ADH10419;
SQ	Sequence 5 AA;	XX	
		DT	11-MAR-2004 (first entry)
	Query Match	DE	HPRG protein H/P rich domain mutated repeat fragment.
	Best Local Similarity	XX	
	Matches	XX	Tpm; tropomyosin; antiangiogenic receptor;
	3; Conservative	KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
	0; Mismatches	KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;
	0; Indels	KW	antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
	0; Gaps	KW	angiogenesis; gene therapy; rabbit.
		OS	Synthetic.
QY	3 PHG 5	XX	
DB	1 PHG 3	XX	
		XX	
		XX	
		XX	
RESULT 14		XX	
AAO15016		XX	
ID	AAO15016 standard; peptide; 5 AA.	XX	
AC	AAO15016;	XX	
XX		XX	
DT	16-AUG-2002 (first entry)	XX	
XX		XX	
DE	Mutant VL CDR3 peptide (C10).	XX	
XX		XX	
KW	Ribosome display system; specific binding pair; sbp; VL CDR3; C10;	XX	
KW	scFv antibody generation; GPI-linked cell surface receptor; mutein.	XX	
XX		XX	
OS	Unidentified.	XX	
OS	Synthetic.	XX	
XX		XX	
PN	WO200175097-A2.	XX	
XX		XX	
PD	11-OCT-2001.	XX	
XX		XX	
PF	26-MAR-2001; 2001WO-GH001319.	XX	
XX		XX	
PR	31-MAR-2000; 2000US-0193802P.	XX	
XX		XX	
PA	(CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.	XX	
XX		XX	
PI	Osbourne J, Holet T;	XX	
XX		XX	
DR	WPI; 2002-010788/01.	XX	
XX		XX	
PT	Obtaining specific binding pair member (I) that binds a complementary	XX	
PT	specific binding pair member of encapsidating specific binding pair	XX	
PT	member/ribosome complexes in a viral coat, comprises the use of a	XX	

CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a mutated HPRG
 CC protein His-Pro (H/P) rich domain repeat fragment.
 XX
 SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db ||||
 3 PHG 5

RESULT 16
 ADH10422
 ID ADH10422 standard; peptide; 5 AA.
 AC ADH10422;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE HPRG protein H/P rich domain mutated repeat fragment.
 XX
 XX Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipsoriatic; dermatological; cardiant; vasotropic; vulnary;
 KW angiogenesis; gene therapy; rabbit.
 XX Synthetic.
 OS
 PN WO2003077872-A2.
 XX
 PD 25-SEP-2003.
 XX
 PF 17-MAR-2003; 2003WO-US008060.
 XX
 PR 15-MAR-2002; 2002US-0364047P.
 XX
 PA (ATTE-) ATTENUON LLC.
 XX
 PI McCrae K, Donate F, Juarez J, Mazar AP;
 XX WPI; 2004-090604/09.
 DR
 XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 XX
 PS Example 5; SEQ ID NO 36; 117pp; English.
 XX The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,

CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a mutated HPRG
 CC protein His-Pro (H/P) rich domain repeat fragment.
 XX
 SQ Sequence 5 AA;

Query Match 91.3%; Score 21; DB 8; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db ||||
 3 PHG 5

RESULT 17
 ADH10417
 ID ADH10417 standard; peptide; 5 AA.
 XX
 AC ADH10417;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Rabbit HPRG protein H/P rich domain repeat fragment.
 XX
 XX Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipsoriatic; dermatological; cardiant; vasotropic; vulnary;
 KW angiogenesis; gene therapy; rabbit.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO2003077872-A2.
 XX
 PD 25-SEP-2003.
 XX
 PF 17-MAR-2003; 2003WO-US008060.
 XX
 PR 15-MAR-2002; 2002US-0364047P.
 XX
 PA (ATTE-) ATTENUON LLC.
 XX
 PI McCrae K, Donate F, Juarez J, Mazar AP;
 XX WPI; 2004-090604/09.
 DR
 XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 XX
 PS Example 5; SEQ ID NO 31; 117pp; English.
 XX The invention relates to an isolated tropomyosin (Tpm)-related
 CC

CC	antiangiogenic receptor polypeptide or peptide, which is a is a fragment	XX	Example 5; SEQ ID NO 29; 117pp; English.
CC	of a full-length native Tpm protein expressed on the surface of	PS	
CC	endothelial cells, or a variant of the fragment. It has a molecular mass	XX	
CC	of about 17 kDa and corresponds in its sequence to, or is a variant of,	CC	The invention relates to an isolated tropomyosin (Tpm)-related
CC	an internal fragment of a native Tpm isoform which is a binding site for	CC	antiangiogenic receptor polypeptide or peptide, which is a is a fragment
CC	antiangiogenic polypeptide agents; The isolated antiangiogenic receptor	CC	of a full-length native Tpm protein expressed on the surface of
CC	polypeptide, peptide or variant has substantially the same biochemical	CC	endothelial cells, or a variant of the fragment. It has a molecular mass
CC	activity of binding to the antiangiogenic polypeptide agents, as does the	CC	of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC	native Tpm internal fragment. The antiangiogenic polypeptide agent that	CC	an internal fragment of a native Tpm isoform which is a binding site for
CC	binds to the isolated polypeptide or peptide is human histidine-proline	CC	antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC	rich glycoprotein (HPRG); rabbit HPRG, a Tpm-binding, antiangiogenic	CC	polypeptide, peptide or variant has substantially the same biochemical
CC	homologue, variant, domain or fragment of human or rabbit HPRG, two chain	CC	activity of binding to the antiangiogenic polypeptide agents, as does the
CC	human kininogen (HK), the D5 domain of HK, or a Tpm-binding, antiangiogenic	CC	native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC	antiangiogenic homologue, variant, domain or fragment of the HK or its D5	CC	binds to the isolated polypeptide or peptide is human histidine-proline
CC	domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,	CC	rich glycoprotein (HPRG); rabbit HPRG, a Tpm-binding, antiangiogenic
CC	antibodies and compositions are useful for inhibiting endothelial cell	CC	homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC	migration, invasion, proliferation or angiogenesis, for inducing	CC	human kininogen (HK), the D5 domain of HK, or a Tpm-binding, antiangiogenic
CC	endothelial cell apoptosis, or for treating tumours or cancer, diabetic	CC	antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC	retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,	CC	domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC	psoriasis, or scleroderma. The antibody may be also used for detecting	CC	antibodies and compositions are useful for inhibiting endothelial cell
CC	the presence of a Tpm polypeptide or peptide in a biological sample, for	CC	migration, invasion, proliferation or angiogenesis, for inducing
CC	promoting wound healing, or for treating diseases or conditions in which	CC	endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC	increased angiogenesis is desired, e.g. coronary artery disease or	CC	retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC	peripheral artery disease. The present sequence represents a repeat	CC	psoriasis, or scleroderma. The antibody may be also used for detecting
CC	fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.	CC	the presence of a Tpm polypeptide or peptide in a biological sample, for
XX		CC	promoting wound healing, or for treating diseases or conditions in which
SQ	Sequence 5 AA;	CC	increased angiogenesis is desired, e.g. coronary artery disease or
		CC	peripheral artery disease. The present sequence represents a repeat
		CC	fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.
		XX	
QY	Query Match	SQ	Sequence 5 AA;
	Best Local Similarity 91.3%; Score 21; DB 8; Length 5;		
	Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
DB	3 PHG 5		
	3 PHG 5		
RESULT 18			
ADH10415			
ID	ADH10415 standard; peptide; 5 AA.		
XX	ADH10415;		
XX			
DT	11-MAR-2004 (first entry)		
DE	Rabbit HPRG protein H/P rich domain repeat fragment.		
XX			
KW	Tpm; tropomyosin; antiangiogenic receptor;		
KW	histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;		
KW	ophthalmological; antiinflammatory; gynaecological; antiarthritic;		
KW	antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;		
KW	angiogenesis; gene therapy; rabbit.		
OS	Oryctolagus cuniculus.		
PN	WO2003077872-A2.		
XX			
PD	25-SEP-2003.		
XX			
PF	17-MAR-2003; 2003WO-US008060.		
XX			
PR	15-MAR-2002; 2002US-0364047P.		
XX			
PA	(ATTE-) ATTENUON LLC.		
XX			
PI	Mccrae K, Donate F, Juarez J, Mazar AP;		
DR	WPI; 2004-090604/09.		
XX			
PT	New tropomyosin-related antiangiogenic receptor polypeptide, useful for		
PT	inhibiting endothelial cell migration, invasion, proliferation or		
PT	angiogenesis, inducing endothelial cell apoptosis, or treating tumors or		
PT	cancer.		

PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.

XX Example 5; SEQ ID NO 30; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

SQ Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PHG 5
|||
Db 3 PHG 5

RESULT 20
ADH10418
XX ADH10418 standard; peptide; 5 AA.
AC ADH10418;
XX
DT 11-MAR-2004 (first entry)
XX
DE HPRG protein H/P rich domain consensus repeat fragment.
XX
XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritis;
KW antipsoriatic; dermatological; cardiac; vasotropic; vulnary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Synthetic.
OS Oryctolagus cuniculus.
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 1 /label= His or Pro
FT Misc-difference 2 /label= His or Pro
FT
XX WO2003077872-A2.
PN
XX

PD 25-SEP-2003.
XX 17-MAR-2003; 2003WO-US008060.
XX 15-MAR-2002; 2002US-0364047P.
XX (ATTE-) ATTENUON LLC.
XX McCrae K, Donate F, Juarez J, Mazar AP;
PI WPI; 2004-090604/09.
XX
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.

XX Example 5; SEQ ID NO 32; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a consensus
CC (H/P) rich domains.

XX Sequence 5 AA;

SQ Query Match 91.3%; Score 21; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 PHG 5
|||
Db 3 PHG 5

RESULT 21
AD081122
XX AD081122 standard; peptide; 5 AA.
AC AD081122;
XX
DT 29-JUL-2004 (first entry)
XX
DE Sheep prion protein motif #3.
XX
XX gene typing; polymorphic microsatellite loci; PML;
KW disease predisposition; microsatellite marker; prion disease;
KW cystic fibrosis; malignant hyperthermia syndrome; metabolic disease;
KW milk protein; hormone; transcription factor; pT7-blue-vector; sheep;
KW microsatellite.

```

XX OS Ovis aries.
XX PN DE10236711-A1.
XX XX
XX PD 26-FEB-2004.
XX PF
XX PF 09-AUG-2002; 2002DE-01036711.
XX PR 09-AUG-2002; 2002DE-01036711.
XX XX (UYHO-) UNIV HOHENHEIM.
XX PI Geldermann H, Preuss S, Han Y;
XX XX WPI; 2004-215730/21.
XX PT Typing genes that contain polymorphic microsatellite loci, useful for
PT identifying predisposition to disease, by amplification and determining
PT length of amplicons.
XX PS Disclosure; Page 9; 64pp; German.
XX CC The invention describes a method of typing (M1) a gene (I) that has one
CC or more polymorphic microsatellite loci (PML). The method comprises: PCR
CC amplification of at least one DNA region of (I) that includes PML, using
CC as template a DNA sample containing at least one segment of (I); and
CC determining the length of the resulting amplicon(s). Also described are:
CC a method of determining (M2) microsatellite markers (MM) for
CC predisposition to a disease, associated with a gene that includes one or
CC more PML; and prediagnosis (M3) of diseases associated with gene that
CC include PML. The method is used to identify microsatellite markers, in a
CC disease-related gene, that are associated with a predisposition to
CC diseases and for prediagnosis of such diseases, especially prion diseases
CC but also cystic fibrosis, malignant hyperthermia syndrome in pigs and
CC metabolic diseases; also to type genes that encode milk proteins,
CC hormones or transcription factors. The method is simpler, quicker and
CC particularly less expensive than known methods based on sequencing. This
CC is the amino acid sequence of a sheep prion protein motif.
XX SQ Sequence 5 AA;
    Query Match 91.3%; Score 21; DB 8; Length 5;
    Best Local Similarity 100.0%; Pred. No. 1.8e+06;
    Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 1 PHG 3

RESULT 22
ADRO3627
ID ADRO3627 standard; peptide; 5 AA.
XX AC
XX AC ADRO3627;
XX DT 21-OCT-2004 (first entry)
XX DE
XX DE E. coli membrane protein ligand #25, SEQ ID 64.
XX KW Ligand; identification; drug target; cancer; cardiovascular disease;
XX KW autoimmune disease; infection; inflammatory disease;
XX KW central nervous system disorder; metabolic disease; endocrine disease;
XX KW membrane protein.
XX OS Synthetic.
XX OS
XX FH Key Location/Qualifiers
XX FT Modified-site 1
XX FT /note= "Xaa = T(Sa), Sialic acid threonine lactam"
XX PN WO2004062553-A2.

XX PD 29-JUL-2004.
XX XX
XX PF 16-JAN-2004; 2004WO-DK000023.
XX PR 16-JAN-2003; 2003US-00346737.
XX PR 19-MAY-2003; 2003DK-00000749.
XX XX (CARL-) CARLSBERG AS.
XX PA
XX PI Hilaire SPM, Yin H, Surve S, Wenckens M;
XX XX WPI; 2004-593147/57.
XX CC Identifying members of protein-ligand binding pair useful as drug target,
XX by incubating immobilized ligand library with differentially labeled
XX protein, detecting ligand-protein pair, identifying protein and ligand of
XX pair.
XX PS Example 33; SEQ ID NO 64; 172pp; English.
XX CC The present invention relates to a method (M1) for identifying specific
XX members of an unknown protein-ligand binding pair. The method involves
XX synthesizing an immobilized ligand library, incubating the library with
XX differentially labelled protein mixtures, detecting the immobilized
XX ligand-protein pair and identifying the protein and ligand of the pair.
XX The identified ligand and protein are specific members of an unknown
XX differential ligand-protein binding pair. (M1) is rapid and efficient,
XX highly accurate and is able to take a large number of unknown ligands
XX and/or unknown proteins. (M1) is used for identifying proteins suitable
XX as drug targets, to identify one or more drugs for treating clinical
XX conditions such as cancer, cardiovascular diseases, autoimmune diseases,
XX infections, inflammatory diseases, central nervous system disorders,
XX metabolic disease and endocrine diseases. The present sequence is a
XX peptide ligand used to illustrate the invention.
XX SQ Sequence 5 AA;
    Query Match 91.3%; Score 21; DB 8; Length 5;
    Best Local Similarity 100.0%; Pred. No. 1.8e+06;
    Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

RESULT 23
AAW75370
ID AAW75370 standard; peptide; 6 AA.
XX AC
XX AC AAW75370;
XX DT 02-FEB-1999 (first entry)
XX DE
XX DE Hexapeptide #10 binds immobilised anti-Ad5 fibre head MAB 7A2.7.
XX KW Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;
XX KW viral surface protein; hexapeptide expression library; adenovirus;
XX KW major histocompatibility complex; MHC; fibronectin; gene therapy;
XX KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.
XX OS Synthetic.
XX OS
XX PN FR2758821-A1.
XX PD 31-JUL-1998.
XX XX
XX PF 30-JAN-1997; 97FR-00001005.
XX PR 30-JAN-1997; 97FR-00001005.
XX XX (CNRS ) CNRS CENT NAT RECH SCI.

```

XX PI Boulanger P, Hong SS, Karayan L;
 XX DR WPI; 1998-416458/36.
 XX
 XX PT Use of peptide(s) for facilitating or modulating attachment of adenovirus
 XX PT to cells - useful for, e.g. treating or preventing infection and
 XX PT improving uptake of gene therapy vectors.
 XX PS Example 3; Fig 2(a); 48pp; French.
 XX CC The invention relates to methods for selecting and identifying a cellular
 CC receptor for a virus, by immobilising, on a support, a monoclonal
 CC antibody targeted to a viral surface protein that determines attachment
 CC of the virus to the receptor. The immobilised antibody is incubated with
 CC a hexapeptide expression library and peptides bound to the immobilised
 CC antibody are eluted by competitive binding with recombinant fragments of
 CC the viral surface protein. In a reverse method, the viral surface protein
 CC is immobilised and incubated with the peptide library. In this case, the
 CC bound peptides are eluted by competitive binding with the monoclonal
 CC antibody. The hexapeptides AAW75361-W75375 represent peptides isolated by
 CC bioassaying the library with an immobilised adenovirus type 5 fibre head
 CC protein and eluted using a monoclonal antibody (MAB) 7A2.7. The methods
 CC are used to identify peptides from MHC Class I and/or type III
 CC fibronectin proteins that allow or facilitate attachment by adenovirus
 CC (Ad) to host cells and/or entry into the cells, and to identify ligands
 CC that modulate Ad infection mediated by these peptides, e.g. to treat or
 CC prevent Ad infections or to facilitate infection by Ad gene therapy
 CC vectors used to treat genetic diseases, acquired immune deficiency
 CC syndrome or cancer.
 XX SQ Sequence 6 AA;
 Query Match 91.3%; Score 21; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 Db |||
 2 PHG 4
 RESULT 24
 AAW75302
 ID AAW75302 standard; peptide; 6 AA.
 XX AC AAW75302;
 XX DT 25-MAR-2003 (revised)
 XX DT 02-FEB-1999 (first entry)
 XX DE Hexapeptide #10 binds immobilised anti-Ads fibre head MAB 7A2.7.
 XX KW Cellular receptor; virus; immobilisation; monoclonal antibody; fibre;
 KW viral surface protein; hexapeptide expression library; adenovirus;
 KW major histocompatibility complex; MHC; fibronectin; gene therapy;
 KW genetic disease; acquired immune deficiency syndrome; AIDS; cancer.
 XX OS Synthetic.
 XX FN FR275885-A1.
 XX PD 31-JUL-1998.
 XX PF 10-JUL-1997; 97FR-00008796.
 XX PR 30-JAN-1997; 97FR-00001005.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX PA Boulanger P, Hong SS, Karayan L;
 XX PI WPI; 1998-416493/36.
 XX DR

XX PT Selection and identification of cellular receptors for viruses - used to
 PT control attachment and entry of adenovirus into cells, e.g. for treating
 PT infection or in gene therapy.
 XX PS Example 3; Fig 2(a); 43pp; French.
 XX CC The invention relates to methods for selecting and identifying a cellular
 CC receptor for a virus, by immobilising, on a support, a monoclonal
 CC antibody targeted to a viral surface protein that determines attachment
 CC of the virus to the receptor. The immobilised antibody is incubated with
 CC a hexapeptide expression library and peptides bound to the immobilised
 CC antibody are eluted by competitive binding with recombinant fragments of
 CC the viral surface protein. In a reverse method, the viral surface protein
 CC is immobilised and incubated with the peptide library. In this case, the
 CC bound peptides are eluted by competitive binding with the monoclonal
 CC antibody. The hexapeptides AAW75293-W75307 represent peptides isolated by
 CC bioassaying the library with an immobilised adenovirus type 5 fibre head
 CC protein and eluted using a monoclonal antibody (MAB) 7A2.7. The methods
 CC are used to identify peptides from MHC Class I and/or type III
 CC fibronectin proteins that allow or facilitate attachment by adenovirus
 CC (Ad) to host cells and/or entry into the cells, and to identify ligands
 CC that modulate Ad infection mediated by these peptides, e.g. to treat or
 CC prevent Ad infections or to facilitate infection by Ad gene therapy
 CC vectors used to treat genetic diseases, acquired immune deficiency
 CC syndrome or cancer. (Updated on 25-MAR-2003 to correct PF field.)
 XX SQ Sequence 6 AA;
 Query Match 91.3%; Score 21; DB 2; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 Db |||
 2 PHG 4
 RESULT 25
 AAB55516
 ID AAB55516 standard; peptide; 6 AA.
 XX AC AAB55516;
 XX DT 07-MAR-2001 (first entry)
 XX DE Human elastase variant segment peptide SEQ ID NO:86.
 XX KW Human; elastase; variant; substrate; mutant; mutagenesis; histidine;
 KW human neutrophil elastase; H43A; cytosolic; proteolysis; ADEPT;
 KW antibody-directed enzyme activated prodrug therapy.
 XX OS Homo sapiens.
 XX FN WO2000068363-A2.
 XX PD 16-NOV-2000.
 XX PF 04-MAY-2000; 2000WO-US006692.
 XX PR 05-MAY-1999; 99US-0132640P.
 XX (GETH) GENENTECH INC.
 XX PI Carter PJ, Dall'acqua W, Rodriques M;
 XX WPI; 2001-007389/01.
 XX PT Elastase variant (H43A) having altered substrate specificity useful for
 PT antibody-directed enzyme activated prodrug therapy.
 XX PS Example 4; Fig 3; 79pp; English.
 XX DR

CC The present invention describes a purified elastase variant (I) with an
 CC amino acid sequence different from that of a precursor elastase, the
 CC difference comprising a substitution of an active site histidine residue
 CC corresponding to residue 43 in human neutrophil elastase with a different
 CC amino acid residue so that (i) has substrate specificity substantially
 CC different from the precursor elastase. (I) has cytotstatic activity, and
 CC can be used in antibody-directed enzyme activated prodrug therapy. The
 CC elastase variant can be used to cleave a particular substrate, especially
 CC those containing histidine residues at the substrate site. Site-specific
 CC proteolysis is useful in therapeutic applications, e.g. for antibody-
 CC directed enzyme activated prodrug therapy (ADEPT). AAC88022, AAC88023 and
 CC AAB5432 to AAB5526 represent sequences used in the exemplification of
 CC the present invention
 XX
 XX
 SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 4; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 2 PHG 4

RESULT 26
 ABG31535
 ID ABG31535 standard; peptide; 6 AA.

AC ABG31535;

DT 05-NOV-2002 (first entry)

DE Camptothecin peptide conjugate #24.

XX Camptothecin; integrin receptor antagonist; cytostatic agent;
 KW tumour cell; metallo matrix protease; MMP; carcinomatous disorder;
 KW integrin alpha.vbeta.3 receptor antagonist; tumour growth inhibitor;
 KW tumour; integrin; camptothecin-bis-trifluoroacetate.

XX Synthetic.

OS Key Location/Qualifiers

FH Modified-site 6

FT /label= OTHER
 FT /note= "Bound to Camptothecin bis-trifluoroacetate at
 FT position 20-0"

XX EP1219305-A1.

XX 03-JUL-2002.

XX 27-DEC-2000; 2000BP-00128401.

XX 27-DEC-2000; 2000BP-00128401.

XX (FARB) BAYER AG.

XX Lerchen H, Baumgarten J, Lockhoff O, Albers M, Schoop A;

XX WPI; 2002-576993/62.

XX New conjugates of integrin receptor antagonist and a cytostatic agent
 XX with specific cleavable linking unit useful in the treatment of cancer.

XX Example; Page 61; 127pp; English.

XX This invention relates to conjugates of integrin receptor antagonist and
 CC a cytostatic agent with cleavable linking unit that are selectively
 CC cleaved by tumour cell metallo matrix proteases (MMPs). The conjugates of
 CC the invention may have cytostatic activity and may be used as an integrin
 CC alpha.vbeta.3 receptor antagonist or a tumour growth inhibitor. The
 CC conjugates of the invention may be used in the production of a medicament

CC for the treatment of carcinomatous disorders. Compounds of the invention
 CC containing the conjugates of the invention exhibit a selective and tumour
 CC -specific action as a result of linkage to alpha.vbeta.3 integrin
 CC antagonists via linking units which can be selectively cleaved by enzymes
 CC such as metallo matrix proteases, i.e. enzymes found in tumour tissue.
 CC The linking units can also maintain the serum stability of the conjugate
 CC of cytostatic and alpha.vbeta.3 integrin antagonist, and at the same
 CC time, show the desired intracellular action within tumour cells as a
 CC result of its specific enzymatic and hydrolytic cleavability with release
 CC of the cytostatic. The present sequence represents a camptothecin-bis-
 CC trifluoroacetate conjugated peptide used in an example of the method of
 CC the invention
 XX
 XX
 SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 1 PHG 3

RESULT 27

AAE26391

ID AAE26391 standard; peptide; 6 AA.

XX AAE26391;

DT 13-DEC-2002 (first entry)

DE Human GPR10 mutant C-terminal peptide, dell16.

XX Human; wakefulness; sleep disorder; prolactin releasing peptide receptor;
 KW PRP; GPR10; therapy; epilepsy; narcolepsy; sleepiness; sleep apnoea;
 KW insomnia; idiopathic hypersomnia; psychogenic hypersomnia; seizure;
 KW anticonvulsant; mutant; mutein.

OS Homo sapiens.

OS Synthetic.

XX US2002037533-A1.

XX 28-MAR-2002.

XX 17-AUG-2001; 2001US-00932161.

XX 28-APR-2000; 2000US-00560915.

XX (CIVE/) CIVELLI O.

XX (LINS/) LIN S.

XX Civelli O, Lin S;

XX WPI; 2002-403931/43.

XX Screening for compounds useful for promoting wakefulness or sleep, and
 XX for treating sleeping disorders, e.g. insomnia, hypersomnia or sleep
 XX apnea, comprises administering a prolactin releasing peptide agonist or
 XX antagonist.

XX Example 2; Page 22; 35pp; English.

XX The present invention relates to a method of screening for compounds for
 CC promoting wakefulness or sleep in a mammal. The method involves
 CC administering a prolactin releasing peptide (PRP) receptor (GPR10)
 CC agonist or antagonist respectively and determining the ability of the
 CC compound to promote wakefulness or sleep. The compounds identified from
 CC the method are used in the therapy of epilepsy and other diseases
 CC associated with absence seizures and in promoting wakefulness and sleep
 CC in individuals having sleep disorders such as insomnia and narcolepsy.
 CC PRP receptor agonists may be used to treat common disorders which lead

CC to sleepiness, e.g. sleep apnoea, narcolepsy, idiopathic hypersomnia and
 CC psychogenic hypersomnia. prp receptor antagonists are useful for
 CC promoting sleep and for treating insomnia such as adjustment sleep
 CC disorder and psychophysiological insomnia. The present sequence is human
 CC GPR10 mutant C-terminal peptide
 XX

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
 Db 1 PHG 3

RESULT 28
 ABP53526
 ID ABP53526 standard; peptide; 6 AA.

XX AC ABP53526;

XX DT 12-DEC-2002 (first entry)

XX DE Camptothecin peptide conjugate II.22.

XX KW Camptothecin peptide conjugate; cytostatic; glycoconjugate; tumour;
 specifically cleavable peptidic linking unit; cancer.

XX OS Synthetic.

XX PH Key Location/Qualifiers

FT Modified-site 1 /note= "N-terminally modified with 20-O-"

FT Modified-site 6 /note= "C-terminally modified with -camptothecin TPA"

PN EP1219634-A1.

XX 03-JUL-2002.

XX PF 27-DEC-2000; 2000EP-00128402.

XX PR 27-DEC-2000; 2000EP-00128402.

XX PA (PABB) BAYER AG.

XX PI Lerchen H, Baumgarten J, Lockhoff O;

XX WPI; 2002-629644/68.

XX Cytostatic-glycoconjugates useful for treating cancer have specifically
 cleavable peptide linking units so as to deliver drug to tumor only.

XX Example; Page 17; 46pp; English.

XX The present invention describes a conjugate (I) and its salts. (I) has
 the formula CT-Li-Sp1-Sp2-K, where: CT = cytotoxic radical or radical of
 a cytostatic derivative optionally also carrying a OH, COOH or NH2 group;
 LI = linker comprising 5 to 8 amino acids each optionally carrying
 protecting groups; Sp1 = absent, CO or CS; Sp2 = optionally substituted
 arylene or alkylene; and K = unsubstituted or regioselectively modified
 carbohydrate radical. Also described: (1) a process for preparation of
 the conjugates; (2) a medicament comprising the conjugate. (I) has
 cytostatic activity. The conjugates are useful for treating cancer. The
 conjugate is cleaved by enzymes found especially in tumour tissue to
 release the cytotoxic drug. As the drug is not active until after
 cleavage it will not cause harm to non-tumour proliferating cells,
 therefore reducing the side-effects associated with the use of such
 drugs. The present sequence represents a camptothecin peptide conjugate,
 CC which is used in an example from the present invention

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 5; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
 Db 1 PHG 3

RESULT 29

ADN12399
 ID ADN12399 standard; peptide; 6 AA.

XX AC ADN12399;

XX DT 17-JUN-2004 (first entry)

XX DE Peptide #1 of the invention.

XX KW ubiquitin-specific protease; neuroplasticity; neurodegeneration.

XX OS Rattus norvegicus.

XX PN WO2003038097-A1.

XX PD 08-MAY-2003.

XX PF 29-OCT-2002; 2002WO-JP011231.

XX PR 29-OCT-2001; 2001JP-00330339.

XX PA (EISA) EISAI CO LTD.

XX PI Suzuki T, Tian QB;

XX DR WPI; 2003-430529/40.

XX Brain-occurring ubiquitin-specific protease and encoded DNA, applicable
 in studying molecular mechanism of expression neuroplasticity and
 pathosis of neurodegeneration.

XX Disclosure; SEQ ID NO 3; 33pp; Japanese.

XX The present invention relates to a protein comprising a 1036 residue
 amino acid sequence having ubiquitin-specific protease activity. The
 protein and its encoded DNA are applicable in studying the molecular
 mechanism of expression neuroplasticity and pathosis of
 neurodegeneration. The present sequence represents the a peptide of the
 invention.

SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 7; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
 Db 4 PHG 6

RESULT 30

ADO37172
 ID ADO37172 standard; peptide; 6 AA.

XX AC ADO37172;

XX DT 29-JUL-2004 (first entry)

XX DE Binding partner polypeptide of the invention SEQ ID NO:232.

KW polypeptide-tagged collection; capture system; tagged polypeptide;
KW pharmaceutical; diagnostic.
XX Synthetic.
OS WO2004039962-A2.
PN 13-MAY-2004.
XX 30-OCT-2003; 2003WO-US034821.
XX 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX (POIN-) POINTILLISTE INC.
XX Ault-Riche D, Atkinson B, Geysen MH;
XX
DR WPI; 2004-376185/35.
XX Evenly distributing tags among members of a starting library, useful in
PT developing pharmaceuticals and diagnostics, comprises dividing the
PT starting library into sub libraries and attaching a tag to members of
PT each sub library.
XX Claim 141; SEQ ID NO 232; 510pp; English.
XX
XX The invention relates to novel methods for producing polypeptide-tagged
CC collections and capture systems containing the tagged polypeptides. The
CC method is useful for evenly distributing tags among members of a starting
CC library. The system, collection, kits and methods are useful in
CC developing pharmaceuticals and diagnostics. The present sequence is used
CC in the exemplification of the invention.
XX
XX Sequence 6 AA;
PI Ault-Riche D, Atkinson B, Geysen MH;
XX
XX Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
DB 1 PHG 3
RESULT 31
ADO37904
ID ADO37904 standard; peptide; 6 AA.
XX ADO37904;
XX 29-JUL-2004 (first entry)
XX Binding partner polypeptide of the invention SEQ ID NO:964.
DE polypeptide-tagged collection; capture system; tagged polypeptide;
KW pharmaceutical; diagnostic.
XX Synthetic.
OS WO2004039962-A2.
PN 13-MAY-2004.
XX 30-OCT-2003; 2003WO-US034821.
XX 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX (POIN-) POINTILLISTE INC.
XX Ault-Riche D, Atkinson B, Geysen MH;
XX

DR WPI; 2004-376185/35.
XX Evenly distributing tags among members of a starting library, useful in
PT developing pharmaceuticals and diagnostics, comprises dividing the
PT starting library into sub libraries and attaching a tag to members of
PT each sub library.
XX Claim 141; SEQ ID NO 964; 510pp; English.
XX The invention relates to novel methods for producing polypeptide-tagged
CC collections and capture systems containing the tagged polypeptides. The
CC method is useful for evenly distributing tags among members of a starting
CC library. The system, collection, kits and methods are useful in
CC developing pharmaceuticals and diagnostics. The present sequence is used
CC in the exemplification of the invention.
XX
XX Sequence 6 AA;
PI Ault-Riche D, Atkinson B, Geysen MH;
XX
XX Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
DB 3 PHG 5
RESULT 32
ADO37705
ID ADO37705 standard; peptide; 6 AA.
XX ADO37705;
XX 29-JUL-2004 (first entry)
XX Binding partner polypeptide of the invention SEQ ID NO:765.
DE polypeptide-tagged collection; capture system; tagged polypeptide;
KW pharmaceutical; diagnostic.
XX Synthetic.
OS WO2004039962-A2.
PN 13-MAY-2004.
XX 30-OCT-2003; 2003WO-US034821.
XX 30-OCT-2002; 2002US-0422923P.
PR 30-OCT-2002; 2002US-0423018P.
XX (POIN-) POINTILLISTE INC.
XX Ault-Riche D, Atkinson B, Geysen MH;
XX
XX WPI; 2004-376185/35.
XX Evenly distributing tags among members of a starting library, useful in
PT developing pharmaceuticals and diagnostics, comprises dividing the
PT starting library into sub libraries and attaching a tag to members of
PT each sub library.
XX Claim 141; SEQ ID NO 765; 510pp; English.
XX The invention relates to novel methods for producing polypeptide-tagged
CC collections and capture systems containing the tagged polypeptides. The
CC method is useful for evenly distributing tags among members of a starting
CC library. The system, collection, kits and methods are useful in
CC developing pharmaceuticals and diagnostics. The present sequence is used
CC in the exemplification of the invention.
XX
XX Sequence 6 AA;
PI Ault-Riche D, Atkinson B, Geysen MH;
XX

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
Db 3 PHG 5

RESULT 33
ADO37180
ID ADO37180 standard; peptide; 6 AA.
XX AC ADO37180;
XX DT 29-JUL-2004 (first entry)
XX DE Binding partner polypeptide of the invention SEQ ID NO:240.
XX KW polypeptide-tagged collection; capture system; tagged polypeptide;
XX KW pharmaceutical; diagnostic.
XX OS Synthetic.
XX PN WO2004039962-A2.
XX PD 13-MAY-2004.
XX PF 30-OCT-2003; 2003WO-US034821.
XX PR 30-OCT-2002; 2002US-0422923P.
XX PR 30-OCT-2002; 2002US-0423018P.
XX PA (POIN-) POINTILLISTE INC.
XX PI Ault-Riche D, Atkinson B, Geysen MH;
XX DR WPI; 2004-376185/35.
XX PT Evenly distributing tags among members of a starting library, useful in
XX PT developing pharmaceuticals and diagnostics, comprises dividing the
XX PT starting library into sub libraries and attaching a tag to members of
XX PT each sub library.
XX PS Claim 141; SEQ ID NO 240; 510pp; English.
XX CC The invention relates to novel methods for producing polypeptide-tagged
XX CC collections and capture systems containing the tagged polypeptides. The
XX CC method is useful for evenly distributing tags among members of a starting
XX CC library. The system, collection, kits and methods are useful in
XX CC developing pharmaceuticals and diagnostics. The present sequence is used
XX CC in the exemplification of the invention.
XX SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
Db 1 PHG 3

RESULT 34
ADO37698
ID ADO37698 standard; peptide; 6 AA.
XX AC ADO37698;
XX DT 29-JUL-2004 (first entry)
XX DE Binding partner polypeptide of the invention SEQ ID NO:758.

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
Db 1 PHG 3

RESULT 35
ADO37897
ID ADO37897 standard; peptide; 6 AA.
XX AC ADO37897;
XX DT 29-JUL-2004 (first entry)
XX DE Binding partner polypeptide of the invention SEQ ID NO:957.
XX KW polypeptide-tagged collection; capture system; tagged polypeptide;
XX KW pharmaceutical; diagnostic.
XX OS Synthetic.
XX PN WO2004039962-A2.
XX PD 13-MAY-2004.
XX PF 30-OCT-2003; 2003WO-US034821.
XX PR 30-OCT-2002; 2002US-0422923P.
XX PR 30-OCT-2002; 2002US-0423018P.
XX PA (POIN-) POINTILLISTE INC.
XX PI Ault-Riche D, Atkinson B, Geysen MH;
XX DR WPI; 2004-376185/35.
XX PT Evenly distributing tags among members of a starting library, useful in
XX PT developing pharmaceuticals and diagnostics, comprises dividing the
XX PT starting library into sub libraries and attaching a tag to members of
XX PT each sub library.
XX PS Claim 141; SEQ ID NO 758; 510pp; English.
XX CC The invention relates to novel methods for producing polypeptide-tagged
XX CC collections and capture systems containing the tagged polypeptides. The
XX CC method is useful for evenly distributing tags among members of a starting
XX CC library. The system, collection, kits and methods are useful in
XX CC developing pharmaceuticals and diagnostics. The present sequence is used
XX CC in the exemplification of the invention.
XX SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
Db 2 PHG 4

RESULT 36
ADO37897
ID ADO37897 standard; peptide; 6 AA.
XX AC ADO37897;
XX DT 29-JUL-2004 (first entry)
XX DE Binding partner polypeptide of the invention SEQ ID NO:957.
XX KW polypeptide-tagged collection; capture system; tagged polypeptide;
XX KW pharmaceutical; diagnostic.
XX OS Synthetic.
XX PN WO2004039962-A2.
XX PD 13-MAY-2004.
XX PF 30-OCT-2003; 2003WO-US034821.
XX PR 30-OCT-2002; 2002US-0422923P.
XX PR 30-OCT-2002; 2002US-0423018P.
XX PA (POIN-) POINTILLISTE INC.
XX PI Ault-Riche D, Atkinson B, Geysen MH;

XX WPI; 2004-376185/35.
 XX Evenly distributing tags among members of a starting library, useful in
 PT developing pharmaceuticals and diagnostics, comprises dividing the
 PT starting library into sub libraries and attaching a tag to members of
 PT each sub library.
 XX
 PS Claim 141; SEQ ID NO 957; 510pp; English.
 XX
 CC The invention relates to novel methods for producing polypeptide-tagged
 CC collections and capture systems containing the tagged polypeptides. The
 CC method is useful for evenly distributing tags among members of a starting
 CC library. The system, collection, kits and methods are useful in
 CC developing pharmaceuticals and diagnostics. The present sequence is used
 CC in the exemplification of the invention.
 XX
 SQ Sequence 6 AA;
 Query Match 91.3%; Score 21; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 Db 2 PHG 4
 RESULT 36
 ADO28227
 ID ADO28227 standard; peptide; 6 AA.
 XX
 AC ADO28227;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Capture system related peptide, SEQ ID 765.
 XX
 KW Capture system.
 XX
 OS Synthetic.
 XX
 PN WO2004042019-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 30-OCT-2003; 2003WO-US034693.
 XX
 PR 30-OCT-2002; 2002US-0422923P.
 PR 30-OCT-2002; 2002US-0423018P.
 XX
 PA (POIN-) POINTILLISTE INC.
 XX
 PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
 XX
 DR WPI; 2004-431543/40.
 XX
 PT Capturing biological particles, by contacting biological particles with
 PT capture system comprising addressed loci, addressed collection of
 PT polypeptide tagged molecules, capture agents, and polypeptide tag to
 PT which capture agent binds.
 XX
 PS Disclosure; SEQ ID NO 765; 505pp; English.
 XX
 CC The present invention relates to a method for the capture and analysis of
 CC biological particle using a capture system. The method is useful for
 CC capturing biological particles such as cells, portions of cells, cell
 CC membranes, viruses, viral capsids, viral particles, bacterial cells,
 CC subcellular compartments, organelles and micelles, prokaryotic cells,
 CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
 CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
 CC vectors or viral capsids with or without packaged nucleic acid, phage,
 CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
 CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
 CC and infected cells, subcellular compartment, organelles, viral particles
 CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
 CC method is also useful for identifying molecules that interact with
 CC infectious agents, for profiling the surface of a biological particles,
 CC for identifying a modulator of an interaction among proteins in the
 CC biological particle, for identifying molecules that modulates the
 CC trafficking, activity or functional or structural property in the
 CC biological particle, and for mapping epitopes of molecules displayed on
 CC the surface of a biological particles. The method is also useful for
 CC sorting biological particles, for identifying a receptor on the surface
 CC of biological particle that transduces a signal from a polypeptide, and
 CC for identifying the molecule that interacts with an apically-localized
 CC molecule on a biological particle. The present sequence was used to
 CC illustrate the invention.
 XX
 SQ Sequence 6 AA;
 Query Match 91.3%; Score 21; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3 PHG 5
 Db 3 PHG 5
 RESULT 37
 ADO28419
 ID ADO28419 standard; peptide; 6 AA.
 XX
 AC ADO28419;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Capture system related peptide, SEQ ID 957.
 XX
 KW Capture system.
 XX
 OS Synthetic.
 XX
 PN WO2004042019-A2.
 XX
 PD 21-MAY-2004.
 XX
 PF 30-OCT-2003; 2003WO-US034693.
 XX
 PR 30-OCT-2002; 2002US-0422923P.
 PR 30-OCT-2002; 2002US-0423018P.
 XX
 PA (POIN-) POINTILLISTE INC.
 XX
 PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
 XX
 DR WPI; 2004-431543/40.
 XX
 PT Capturing biological particles, by contacting biological particles with
 PT capture system comprising addressed loci, addressed collection of
 PT polypeptide tagged molecules, capture agents, and polypeptide tag to
 PT which capture agent binds.
 XX
 PS Disclosure; SEQ ID NO 957; 505pp; English.
 XX
 CC The present invention relates to a method for the capture and analysis of
 CC biological particle using a capture system. The method is useful for
 CC capturing biological particles such as cells, portions of cells, cell
 CC membranes, viruses, viral capsids, viral particles, bacterial cells,
 CC subcellular compartments, organelles and micelles, prokaryotic cells,
 CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
 CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
 CC vectors or viral capsids with or without packaged nucleic acid, phage,
 CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
 CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
 CC and infected cells, subcellular compartment, organelles, viral particles
 CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
 CC method is also useful for identifying molecules that interact with
 CC infectious agents, for profiling the surface of a biological particles,
 CC for identifying a modulator of an interaction among proteins in the
 CC biological particle, for identifying molecules that modulates the
 CC trafficking, activity or functional or structural property in the
 CC biological particle, and for mapping epitopes of molecules displayed on
 CC the surface of a biological particles. The method is also useful for
 CC sorting biological particles, for identifying a receptor on the surface
 CC of biological particle that transduces a signal from a polypeptide, and
 CC for identifying the molecule that interacts with an apically-localized
 CC molecule on a biological particle. The present sequence was used to
 CC illustrate the invention.

XX
 SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
 Db 2 PHG 4

RESULT 38
 ADO28426
 ID ADO28426 standard; peptide; 6 AA.

XX ADO28426;

XX 12-AUG-2004 (first entry)

DT Capture system related peptide, SEQ ID 964.

DE Capture system.

KW Synthetic.

OS WO2004042019-A2.

PN 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034693.

XX 30-OCT-2002; 2002US-0422923P.

PR 30-OCT-2002; 2002US-0423018P.

XX (POIN-) POINTILLISTE INC.

XX Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;

XX WPI; 2004-431543/40.

XX Capturing biological particles, by contacting biological particles with
 PT capture system comprising addressed loci, addressed collection of
 PT polypeptide tagged molecules, capture agents, and polypeptide tag to
 PT which capture agent binds.

PS Disclosure; SEQ ID NO 964; 505pp; English.

XX The present invention relates to a method for the capture and analysis of
 CC biological particle using a capture system. The method is useful for
 CC capturing biological particles such as cells, portions of cells, cell
 CC membranes, viruses, viral capsids, viral particles, bacterial cells,
 CC subcellular compartments, organelles and micelles, prokaryotic cells,
 CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
 CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
 CC vectors or viral capsids with or without packaged nucleic acid, phage,
 CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
 CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
 CC and infected cells, subcellular compartment, organelles, viral particles
 CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
 CC method is also useful for identifying molecules that interact with
 CC infectious agents, for profiling the surface of a biological particles,
 CC for identifying a modulator of an interaction among proteins in the
 CC biological particle, for identifying molecules that modulates the
 CC trafficking, activity or functional or structural property in the
 CC biological particle, and for mapping epitopes of molecules displayed on
 CC the surface of a biological particles. The method is also useful for
 CC sorting biological particles, for identifying a receptor on the surface
 CC of biological particle that transduces a signal from a polypeptide, and
 CC for identifying the molecule that interacts with an apically-localized
 CC molecule on a biological particle. The present sequence was used to
 CC illustrate the invention.

XX
 SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 |||
 Db 3 PHG 5

RESULT 39
 ADO27694
 ID ADO27694 standard; peptide; 6 AA.

XX ADO27694;

XX 12-AUG-2004 (first entry)

DT Capture system related peptide, SEQ ID 232.

DE Capture system.

KW Synthetic.

OS WO2004042019-A2.

PN 21-MAY-2004.

XX 30-OCT-2003; 2003WO-US034693.

XX 30-OCT-2002; 2002US-0422923P.

PR 30-OCT-2002; 2002US-0423018P.

XX (POIN-) POINTILLISTE INC.

XX Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;

XX WPI; 2004-431543/40.

XX Capturing biological particles, by contacting biological particles with
 PT capture system comprising addressed loci, addressed collection of
 PT polypeptide tagged molecules, capture agents, and polypeptide tag to
 PT which capture agent binds.

PS Disclosure; SEQ ID NO 232; 505pp; English.

XX The present invention relates to a method for the capture and analysis of
 CC biological particle using a capture system. The method is useful for
 CC capturing biological particles such as cells, portions of cells, cell
 CC membranes, viruses, viral capsids, viral particles, bacterial cells,
 CC subcellular compartments, organelles and micelles, prokaryotic cells,
 CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
 CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
 CC vectors or viral capsids with or without packaged nucleic acid, phage,
 CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.

XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
1 PHG 3

RESULT 40
ADO27702
ID ADO27702 standard; peptide; 6 AA.
XX
AC ADO27702;
XX
DT 12-AUG-2004 (first entry)
XX
DE Capture system related peptide, SEQ ID 240.
XX
KW Capture system.
XX
OS Synthetic.
XX
FN WO2004042019-A2.
XX
PD 21-MAY-2004.
XX
PF 30-OCT-2003; 2003WO-US034693.
XX
PR 30-OCT-2002; 2002US-0422923P.
XX
PR 30-OCT-2002; 2002US-0423018P.
XX
PA (POIN-) POINTILLISTE INC.
XX
PI Ault-Riche D, Atkinson B, Jesaitis L, Kumble KD, Sperinde G;
XX
DR WPI; 2004-431543/40.
XX
PT Capturing biological particles, by contacting biological particles with
PT capture system comprising addressed loci, addressed collection of
PT polypeptide tagged molecules, capture agents, and polypeptide tag to
PT which capture agent binds.
XX
PS Disclosure; SEQ ID NO 240; 505pp; English.
XX
CC The present invention relates to a method for the capture and analysis of
CC biological particle using a capture system. The method is useful for
CC capturing biological particles such as cells, portions of cells, cell
CC membranes, viruses, viral capsids, viral particles, bacterial cells,
CC subcellular compartments, organelles and micelles, prokaryotic cells,
CC eukaryotic cells, intracellular particles, nuclei, cell membranes, cell
CC membrane fragments, nuclear membranes, nuclear membranes fragments, viral
CC vectors or viral capsids with or without packaged nucleic acid, phage,
CC phage vectors, phage capsids with or without encapsulated nucleotide

CC acid, liposomes and other micellar agents. The biological particles are
CC cells chosen from immune cells, neurons, cancer cells, bacterial cells
CC and infected cells, subcellular compartment, organelles, viral particles
CC or pathogens. The cells are dendritic cells, T cells, or B cells. The
CC method is also useful for identifying molecules that interact with
CC infectious agents, for profiling the surface of a biological particles,
CC for identifying a modulator of an interaction among proteins in the
CC biological particle, for identifying molecules that modulates the
CC trafficking, activity or functional or structural property in the
CC biological particle, and for mapping epitopes of molecules displayed on
CC the surface of a biological particles. The method is also useful for
CC sorting biological particles, for identifying a receptor on the surface
CC of biological particle that transduces a signal from a polypeptide, and
CC for identifying the molecule that interacts with an apically-localized
CC molecule on a biological particle. The present sequence was used to
CC illustrate the invention.

XX
SQ Sequence 6 AA;

Query Match 91.3%; Score 21; DB 8; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
1 PHG 3

Search completed: June 15, 2005, 14:15:45
JOB time : 94 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-7

Perfect score: 23

Sequence: 1 XPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/iaa/5A_COMB.pep.*
- 2: /cgn2_6/prodata/1/iaa/5B_COMB.pep.*
- 3: /cgn2_6/prodata/1/iaa/6A_COMB.pep.*
- 4: /cgn2_6/prodata/1/iaa/6B_COMB.pep.*
- 5: /cgn2_6/prodata/1/iaa/PCFUS_COMB.pep.*
- 6: /cgn2_6/prodata/1/iaa/backfilee1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	91.3	5	1	US-07-729-099-12
2	21	91.3	5	1	US-08-257-392-12
3	21	91.3	5	3	US-08-770-035-12
4	21	91.3	5	4	US-10-122-246A-64
5	21	91.3	5	6	5217869-59
6	21	91.3	5	6	5217869-59
7	21	91.3	6	3	US-09-095-407-3
8	21	91.3	6	3	US-09-560-915-5
9	21	91.3	6	4	US-09-155-613A-71
10	21	91.3	6	5	PCT-US95-04567-11
11	21	91.3	7	3	US-08-776-265-12
12	21	91.3	7	4	US-09-398-184-12
13	21	91.3	8	3	US-08-405-647B-13
14	21	91.3	8	3	US-08-985-499-13
15	21	91.3	8	3	US-09-125-099-2
16	21	91.3	8	4	US-09-239-043D-1149
17	21	91.3	8	4	US-09-239-043D-1209
18	21	91.3	8	4	US-09-112-956-1
19	21	91.3	8	5	PCT-US96-03180-13
20	21	91.3	9	3	US-09-510-738A-102
21	21	91.3	9	3	US-09-510-738A-124
22	21	91.3	9	3	US-09-510-738A-125
23	21	91.3	9	3	US-09-510-738A-140
24	21	91.3	9	3	US-09-510-738A-144
25	21	91.3	9	4	US-09-861-966-102
26	21	91.3	9	4	US-09-861-966-124
27	21	91.3	9	4	US-09-861-966-125

28	21	91.3	9	4	US-09-861-966-140	Sequence 140, App
29	21	91.3	9	4	US-09-861-966-144	Sequence 144, App
30	21	91.3	9	4	US-09-239-043D-1128	Sequence 1128, Ap
31	21	91.3	9	4	US-09-239-043D-1667	Sequence 1667, Ap
32	21	91.3	9	4	US-09-870-089B-13	Sequence 13, Appl
33	21	91.3	9	4	US-09-919-048-102	Sequence 102, App
34	21	91.3	9	4	US-09-919-048-124	Sequence 124, App
35	21	91.3	9	4	US-09-919-048-125	Sequence 125, App
36	21	91.3	9	4	US-09-919-048-140	Sequence 140, App
37	21	91.3	9	4	US-09-919-048-144	Sequence 144, App
38	21	91.3	10	1	US-08-049-511-10	Sequence 10, Appl
39	21	91.3	10	1	US-08-049-511-11	Sequence 11, Appl
40	21	91.3	10	1	US-08-049-511-13	Sequence 13, Appl
41	21	91.3	10	2	US-08-556-597-92	Sequence 92, Appl
42	21	91.3	10	2	US-08-318-856A-56	Sequence 56, Appl
43	21	91.3	10	3	US-09-043-930-14	Sequence 14, Appl
44	21	91.3	10	4	US-09-258-689-3	Sequence 3, Appli
45	21	91.3	10	4	US-09-195-379-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-07-729-099-12
; Sequence 12, Application US/07729099
; Patent No. 5403581
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Pasamontes, Luis
; TITLE OF INVENTION: Coccidiosis Vaccines
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07110
; COMPUTER READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/729,099
; FILING DATE: 19910712
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Bimeria tenella
; US-07-729-099-12

Query Match 91.3%; Score 21; DB 1; Length 5;
Best Local Similarity 100.0%; Pred.No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||

Db 3 PHG 5

RESULT 2

US-08-257-392-12
; Sequence 12, Application US/08257392
; Patent No. 5688513
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Pasamontes, Luis
; TITLE OF INVENTION: Coccidiosis Vaccines
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey
; COUNTRY: U.S.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,392
; FILING DATE: 09-JUN-1994
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/729,099
; FILING DATE: 12-JUL-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Eimeria tenella
; US-08-257-392-12

Query Match 91.3%; Score 21; DB 1; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 3 PHG 5

RESULT 3

US-08-770-035-12
; Sequence 12, Application US/08770035
; Patent No. 6008342
; GENERAL INFORMATION:
; APPLICANT: Binger, Mary-Helen
; APPLICANT: Pasamontes, Luis
; TITLE OF INVENTION: Coccidiosis Vaccines
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Hoffmann-La Roche Inc.
; STREET: 340 Kingsland Street
; CITY: Nutley
; STATE: New Jersey

COUNTRY: U.S.
; ZIP: 07110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/770,035
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/729,099
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Roseman, Catherine R
; REGISTRATION NUMBER: 34,240
; REFERENCE/DOCKET NUMBER: 8514
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (201) 235-6208
; TELEFAX: (201) 235-3500
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: YES
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Eimeria tenella
; US-08-770-035-12

Query Match 91.3%; Score 21; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 3 PHG 5

RESULT 4

US-10-122-246A-64
; Sequence 64, Application US/10122246A
; Patent No. 6767897
; GENERAL INFORMATION:
; APPLICANT: ABAJIAN, Henry B
; APPLICANT: HLAVKA, Joseph J
; APPLICANT: FEIGNER, John P
; TITLE OF INVENTION: NEW THERAPEUTIC USES OF TRI-, TETRA-, PENTA-, AND POLYPEPTIDES
; FILE REFERENCE: D4850-00021
; CURRENT APPLICATION NUMBER: US/10/122,246A
; CURRENT FILING DATE: 2002-04-11
; PRIOR APPLICATION NUMBER: US 09/625,103
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 08/962,962
; PRIOR FILING DATE: 1997-11-04
; PRIOR APPLICATION NUMBER: US 08/432,651
; PRIOR FILING DATE: 1995-05-02
; PRIOR APPLICATION NUMBER: US 08/238,089
; PRIOR FILING DATE: 1994-05-04
; NUMBER OF SEQ ID NOS: 110
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 64
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; FEATURE:
; NAME/KEY: MOD_RES
; LOCATION: (1)..(1)

```
; OTHER INFORMATION: 4-fluorophenylalanine
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (2)-(2)
; OTHER INFORMATION: 4HYP
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (5)-(5)
; OTHER INFORMATION: AMIDATION
US-10-122-246A-64

Query Match          91.3%; Score 21; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 5
5217869-59
; Patent No. 5217869
; APPLICANT: KAUVAR, LAWRENCE M.
; TITLE OF INVENTION: METHOD TO PRODUCE IMMUNODIAGNOSTIC
; REAGENTS
; NUMBER OF SEQUENCES: 121
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/255,906
; FILING DATE: 11-OCT-1988
; SEQ ID NO: 59;
; LENGTH: 5
5217869-59

Query Match          91.3%; Score 21; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 6
5217869-59
; Patent No. 5217869
; APPLICANT: KAUVAR, LAWRENCE M.
; TITLE OF INVENTION: METHOD TO PRODUCE IMMUNODIAGNOSTIC
; REAGENTS
; NUMBER OF SEQUENCES: 121
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/255,906
; FILING DATE: 11-OCT-1988
; SEQ ID NO: 59;
; LENGTH: 5
5217869-59

Query Match          91.3%; Score 21; DB 6; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 7
US-09-095-407-3
; Sequence 3, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; OTHER INFORMATION: Yip, Tai-Tung
; US-09-095-407-3

TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
;
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-095-407-3

Query Match          91.3%; Score 21; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      4 PHG 6

RESULT 8
US-09-560-915-5
; Sequence 5, Application US/09560915
; Patent No. 6383764
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; TITLE OF INVENTION: Relating To Prolactin Releasing Peptide (PrRP)
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/09/560,915
; CURRENT FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human GPR10 variant
; US-09-560-915-5
```

```
Query Match          91.3%; Score 21; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 9
US-09-155-613A-71
; Sequence 71, Application US/09155613A
; Patent No. 6420120
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/09/155,613A
; CURRENT FILING DATE: 1998-09-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotopoe
US-09-155-613A-71

Query Match          91.3%; Score 21; DB 4; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 10
PCT-US95-04567-11
; Sequence 11, Application PC/TUS9504567
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: DIHYDROPYRIMIDINE DEHYDROGENASE
; TITLE OF INVENTION: COMPOSITIONS AND METHODS OF USE
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P.O. Box 4433
; CITY: Houston
; STATE: Texas
; COUNTRY: United States of America
; ZIP: 77210
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS/ASCII
; SOFTWARE: #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/04567
; FILING DATE: CONCURRENTLY HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: US 08/227,357
; FILING DATE: 13-APR-1994
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Wilson, Mark B.
; REGISTRATION NUMBER: 37,259
; REFERENCE/DOCKET NUMBER: UOAB025P--
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (512) 418-3000
; TELEFAX: (713) 789-2679
; TELEX: 79-0924
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US95-04567-11

Query Match          91.3%; Score 21; DB 5; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 11
US-08-776-265-12
; Sequence 12, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; NUCLEOTIDE SEQUENCES AND USES THEREOF
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
```

US-08-776-265-12

Query Match 91.3%; Score 21; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 3 PHG 5

RESULT 12

US-09-398-184-12
; Sequence 12, Application US/09398184
; Patent No. 6649394

GENERAL INFORMATION:

APPLICANT: BLANCHE, Francis
CAMERON, Beatrice
CROUZET, Joel
FAMECHON, Alain
FERRERO, Lucia

TITLE OF INVENTION: No. 6649394el Topoisomerase IV, Corresponding
Nucleotide Sequences and Uses Thereof

NUMBER OF SEQUENCES: 17

CORRESPONDENCE ADDRESS:

ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &

STREET: 1300 I. Street, N.W., Suite 700

CITY: Washington

STATE: D.C.

COUNTRY: USA

ZIP: 20005-3315

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/398,184

FILING DATE: 17-Sep-1999

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/776,265

FILING DATE: 24-JAN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Einaudi, Carol P.

REGISTRATION NUMBER: 32,220

REFERENCE/DOCKET NUMBER: 03806.0394-00000

TELECOMMUNICATION INFORMATION:

TELEPHONE: (202) 408-4000

TELEFAX: (202) 408-4444

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 7 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

SEQUENCE DESCRIPTION: SEQ ID NO: 12:

US-09-398-184-12

Query Match 91.3%; Score 21; DB 4; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 3 PHG 5

RESULT 13

US-08-405-647B-13

; Sequence 13, Application US/08405647B

; Patent No. 6124262
; GENERAL INFORMATION:
; APPLICANT: Sherman, Irwin W.
; APPLICANT: Crandall, Ian E.
; APPLICANT: Sholet, Stephen B.
; APPLICANT: Thevenin, Bernard Jean-Marie
; TITLE OF INVENTION: Compositions and Methods for Reducing
; TITLE OF INVENTION: Adhesiveness of Defective Red Blood Cells
; NUMBER OF SEQUENCES: 50
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/405,647B

FILING DATE: 17-MAR-1995

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Weber, Kenneth A.

REGISTRATION NUMBER: 31,677

REFERENCE/DOCKET NUMBER: 02307B-068700US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 8 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-405-647B-13

Query Match 91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 6 PHG 8

RESULT 14

US-08-985-499-13

; Sequence 13, Application US/08985499

; Patent No. 6191103

; GENERAL INFORMATION:

APPLICANT: Sholet, Stephen B.

APPLICANT: Sherman, Irwin

APPLICANT: von Andrian, Ulrich

TITLE OF INVENTION: Methods for Enhancing Thrombolysis in a

TITLE OF INVENTION: Mammal

NUMBER OF SEQUENCES: 45

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP

STREET: Two Embarcadero Center, Eighth Floor

CITY: San Francisco

STATE: California

COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

```
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/985,499
/ FILING DATE: 05-DEC-1997
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Hyman, Laurence J.
/ REGISTRATION NUMBER: 35,551
/ REFERENCE/DOCKET NUMBER: 02307E-084500US
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (415) 576-0200
/ TELEFAX: (415) 576-0300
/ INFORMATION FOR SEQ ID NO: 13:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 8 amino acids
/ TYPE: amino acid
/ STRANDEDNESS:
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ US-08-985-499-13

Query Match          91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      6 PHG 8

RESULT 15
US-09-125-099-2
/ Sequence 2, Application US/09125099A
/ Patent No. 6207150
/ GENERAL INFORMATION:
/ APPLICANT: CROUZET, Joel
/ APPLICANT: BLANCHE, Francis
/ APPLICANT: COUDER, Michel
/ APPLICANT: CAMERON, Beatrice
/ TITLE OF INVENTION: VARIANTS OF THYMIDINE KINASE, RELATED NUCLEIC ACIDS
/ FILE REFERENCE: ST96010-US
/ CURRENT APPLICATION NUMBER: US/09/125,099A
/ CURRENT FILING DATE: 1998-08-06
/ EARLIER APPLICATION NUMBER: PCT/FR97/00193
/ EARLIER FILING DATE: 1997-01-31
/ EARLIER APPLICATION NUMBER: FR96/01603
/ EARLIER FILING DATE: 1996-02-09
/ EARLIER APPLICATION NUMBER: FR96/09709
/ EARLIER FILING DATE: 1996-08-01
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 8
/ TYPE: PRT
/ ORGANISM: herpes simplex virus 7
/ US-09-125-099-2

Query Match          91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      2 PHG 4

RESULT 16
US-09-239-043D-1149
/ Sequence 1149, Application US/09239043D
/ Patent No. 6689363
/ GENERAL INFORMATION:
/ APPLICANT: Sette, Alessandro
/ APPLICANT: Sidney, John

Query Match          91.3%; Score 21; DB 3; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      3 PHG 5

RESULT 17
US-09-239-043D-1209
/ Sequence 1209, Application US/09239043D
/ Patent No. 6689363
/ GENERAL INFORMATION:
/ APPLICANT: Sette, Alessandro
/ APPLICANT: Sidney, John
/ APPLICANT: Southwood, Scott
/ APPLICANT: Vitiello, Maria A.
/ APPLICANT: Livingston, Brian D.
/ APPLICANT: Celis, Esteban
/ APPLICANT: Kubo, Ralph T.
/ APPLICANT: Grey, Howard M.
/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 09/189,702
/ PRIOR FILING DATE: 1998-11-10
```

```
/ APPLICANT: Southwood, Scott
/ APPLICANT: Vitiello, Maria A.
/ APPLICANT: Livingston, Brian D.
/ APPLICANT: Celis, Esteban
/ APPLICANT: Kubo, Ralph T.
/ APPLICANT: Grey, Howard M.
/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 09/189,702
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: US 08/978,291
/ PRIOR FILING DATE: 1997-11-25
/ PRIOR APPLICATION NUMBER: US 08/820,360
/ PRIOR FILING DATE: 1997-03-12
/ PRIOR APPLICATION NUMBER: US 60/013,363
/ PRIOR FILING DATE: 1996-03-13
/ PRIOR APPLICATION NUMBER: US 08/461,603
/ PRIOR FILING DATE: 1995-06-05
/ PRIOR APPLICATION NUMBER: US 08/347,610
/ PRIOR FILING DATE: 1994-12-01
/ PRIOR APPLICATION NUMBER: US 08/344,824
/ PRIOR FILING DATE: 1994-11-23
/ PRIOR APPLICATION NUMBER: US 08/278,634
/ PRIOR FILING DATE: 1994-07-21
/ PRIOR APPLICATION NUMBER: US 08/205,713
/ PRIOR FILING DATE: 1994-03-04
/ PRIOR APPLICATION NUMBER: US 08/197,484
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 2579
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1149
/ LENGTH: 8
/ TYPE: PRT
/ ORGANISM: Orthohepadnaviridae hepatitis B virus
/ US-09-239-043D-1149

Query Match          91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      3 PHG 5

RESULT 17
US-09-239-043D-1209
/ Sequence 1209, Application US/09239043D
/ Patent No. 6689363
/ GENERAL INFORMATION:
/ APPLICANT: Sette, Alessandro
/ APPLICANT: Sidney, John
/ APPLICANT: Southwood, Scott
/ APPLICANT: Vitiello, Maria A.
/ APPLICANT: Livingston, Brian D.
/ APPLICANT: Celis, Esteban
/ APPLICANT: Kubo, Ralph T.
/ APPLICANT: Grey, Howard M.
/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 09/189,702
/ PRIOR FILING DATE: 1998-11-10
```

PRIOR APPLICATION NUMBER: US 08/978,291
PRIOR FILING DATE: 1997-11-25
PRIOR APPLICATION NUMBER: US 08/820,360
PRIOR FILING DATE: 1997-03-12
PRIOR APPLICATION NUMBER: US 60/013,363
PRIOR FILING DATE: 1996-03-13
PRIOR APPLICATION NUMBER: US 08/461,603
PRIOR FILING DATE: 1995-06-05
PRIOR APPLICATION NUMBER: US 08/347,610
PRIOR FILING DATE: 1994-12-01
PRIOR APPLICATION NUMBER: US 08/344,824
PRIOR FILING DATE: 1994-11-23
PRIOR APPLICATION NUMBER: US 08/278,634
PRIOR FILING DATE: 1994-07-21
PRIOR APPLICATION NUMBER: US 08/205,713
PRIOR FILING DATE: 1994-03-04
PRIOR APPLICATION NUMBER: US 08/197,484
PRIOR FILING DATE: 1994-02-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 2579
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 1209
LENGTH: 8
TYPE: PRT
ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1209

Query Match 91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 1 PHG 3

RESULT 18
US-09-112-956-1
Sequence 1, Application US/09112956
Patent No. 6750025
GENERAL INFORMATION:
APPLICANT: Hammond, David
APPLICANT: Medina, Emma
TITLE OF INVENTION: A NOVEL METHOD OF DETECTING AND ISOLATING PRION
TITLE OF INVENTION: PROTEIN AND VARIANTS THEREOF, AND NOVEL METHODS OF
DIAGNOSING AND TREATING PRION DISEASES
FILE REFERENCE: 92053/26
CURRENT APPLICATION NUMBER: US/09/112,956
CURRENT FILING DATE: 1998-07-09
NUMBER OF SEQ ID NOS: 6
SOFTWARE: Patentin version 3.0
SEQ ID NO 1
LENGTH: 8
TYPE: PRT
ORGANISM: homo sapiens
US-09-112-956-1

Query Match 91.3%; Score 21; DB 4; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 5 PHG 7

RESULT 19
PCT-US96-03180-13
Sequence 13, Application PC/TUS9603180
GENERAL INFORMATION:
APPLICANT: The Regents of the University
APPLICANT: of California
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR REDUCING

TITLE OF INVENTION: ADHESIVENESS OF DEFECTIVE RED BLOOD CELLS
NUMBER OF SEQUENCES: 43
CORRESPONDENCE ADDRESS:
ADDRESSEE: Robbins, Berliner & Carson
STREET: 201 North Figueroa Street, Suite 500
CITY: Los Angeles
STATE: California
COUNTRY: USA
ZIP: 90012-2628
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/03180
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Berliner, Robert
REGISTRATION NUMBER: 20,121
REFERENCE/DOCKET NUMBER: 5555-370
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 977-1001
TELEFAX: (213) 977-1003
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 8 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US96-03180-13

Query Match 91.3%; Score 21; DB 5; Length 8;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 6 PHG 8

RESULT 20
US-09-510-738A-102
Sequence 102, Application US/09510738A
Patent No. 6268165
GENERAL INFORMATION:
APPLICANT: O'Brien, Timothy J.
TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
Ovarian Cancer
FILE REFERENCE: D6223CIP-A
CURRENT APPLICATION NUMBER: US/09/510,738A
CURRENT FILING DATE: 2000-02-22
PRIOR APPLICATION NUMBER: 09/039,211
PRIOR FILING DATE: 03-14-1998
NUMBER OF SEQ ID NOS: 188
SEQ ID NO 102
LENGTH: 9
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-102

Query Match 91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 6 PHG 8

```
RESULT 21
US-09-510-738A-124
; Sequence 124, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-124

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 22
US-09-510-738A-125
; Sequence 125, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-510-738A-125

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 23
US-09-510-738A-140
; Sequence 140, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-510-738A-140

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 24
US-09-510-738A-144
; Sequence 144, Application US/09510738A
; Patent No. 6268165
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A
; CURRENT APPLICATION NUMBER: US/09/510,738A
; CURRENT FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: 09/039,211
; PRIOR FILING DATE: 03-14-1998
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-510-738A-144

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      6 PHG 8

RESULT 25
US-09-861-966-102
; Sequence 102, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; TITLE OF INVENTION: Ovarian Cancer
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 102
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-102

Query Match          91.3%; Score 21; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      4 PHG 6
```

```
Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 26
US-09-861-966-124
; Sequence 124, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-124

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 27
US-09-861-966-125
; Sequence 125, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-861-966-125

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      4 PHG 6

RESULT 28
US-09-861-966-140
; Sequence 140, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-861-966-140

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      6 PHG 8

RESULT 29
US-09-861-966-144
; Sequence 144, Application US/09861966
; Patent No. 6518028
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; TITLE OF INVENTION: Compositions and Methods for the Early Diagnosis of
; FILE REFERENCE: D6223CIP-A/Div
; CURRENT APPLICATION NUMBER: US/09/861,966
; CURRENT FILING DATE: 2001-05-21
; PRIOR APPLICATION NUMBER: 09/510,738
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-861-966-144

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      4 PHG 6

RESULT 30
US-09-239-043D-1128
; Sequence 1128, Application US/09239043D
; Patent No. 6689363
; GENERAL INFORMATION:
; APPLICANT: Sette, Alessandro
; APPLICANT: Sidney, John
; APPLICANT: Southwood, Scott
; APPLICANT: Vitello, Maria A.
; APPLICANT: Livingston, Brian D.
; APPLICANT: Celis, Esteban
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
```

```
/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE OF INVENTION: Using Peptide and Nucleic Acid Compositions
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 09/189,702
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: US 08/978,291
/ PRIOR FILING DATE: 1997-11-25
/ PRIOR APPLICATION NUMBER: US 08/820,360
/ PRIOR FILING DATE: 1997-03-12
/ PRIOR APPLICATION NUMBER: US 60/013,363
/ PRIOR FILING DATE: 1996-03-13
/ PRIOR APPLICATION NUMBER: US 08/461,603
/ PRIOR FILING DATE: 1995-06-05
/ PRIOR APPLICATION NUMBER: US 08/347,610
/ PRIOR FILING DATE: 1994-12-01
/ PRIOR APPLICATION NUMBER: US 08/344,824
/ PRIOR FILING DATE: 1994-11-23
/ PRIOR APPLICATION NUMBER: US 08/278,634
/ PRIOR FILING DATE: 1994-07-21
/ PRIOR APPLICATION NUMBER: US 08/205,713
/ PRIOR FILING DATE: 1994-03-04
/ PRIOR APPLICATION NUMBER: US 08/197,484
/ PRIOR FILING DATE: 1994-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 2579
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 1128
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1128

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 31
US-09-239-043D-1667
/ Sequence 1667, Application US/09239043D
/ Patent No. 6689363
/ GENERAL INFORMATION:
/ APPLICANT: Sette, Alessandro
/ APPLICANT: Sidney, John
/ APPLICANT: Southwood, Scott
/ APPLICANT: Vitello, Maria A.
/ APPLICANT: Livingston, Brian D.
/ APPLICANT: Celis, Esteban
/ APPLICANT: Kubo, Ralph T.
/ APPLICANT: Grey, Howard M.
/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE OF INVENTION: Using Peptide and Nucleic Acid Compositions
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 09/189,702
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: US 08/978,291
/ PRIOR FILING DATE: 1997-11-25
/ PRIOR APPLICATION NUMBER: US 08/820,360
/ PRIOR FILING DATE: 1997-03-12
/ PRIOR APPLICATION NUMBER: US 60/013,363
/ PRIOR FILING DATE: 1996-03-13

/ APPLICANT: Chesnut, Robert
/ APPLICANT: Epimmune Inc.
/ TITLE OF INVENTION: Inducing Cellular Immune Responses to Hepatitis B Virus
/ FILE OF INVENTION: Using Peptide and Nucleic Acid Compositions
/ FILE REFERENCE: 2060.0060007
/ CURRENT APPLICATION NUMBER: US/09/239,043D
/ CURRENT FILING DATE: 1999-01-27
/ PRIOR APPLICATION NUMBER: US 08/978,291
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: US 08/820,360
/ PRIOR FILING DATE: 1997-03-12
/ PRIOR APPLICATION NUMBER: US 60/013,363
/ PRIOR FILING DATE: 1996-03-13
```

```
/ PRIOR APPLICATION NUMBER: US 08/461,603
/ PRIOR FILING DATE: 1995-06-05
/ PRIOR APPLICATION NUMBER: US 08/347,610
/ PRIOR FILING DATE: 1994-12-01
/ PRIOR APPLICATION NUMBER: US 08/344,824
/ PRIOR FILING DATE: 1994-11-23
/ PRIOR APPLICATION NUMBER: US 08/278,634
/ PRIOR FILING DATE: 1994-07-21
/ PRIOR APPLICATION NUMBER: US 08/205,713
/ PRIOR FILING DATE: 1994-03-04
/ PRIOR APPLICATION NUMBER: US 08/197,484
/ PRIOR FILING DATE: 1994-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 2579
/ SOFTWARE: FastSEQ for Windows Version 4.0
/ SEQ ID NO 1667
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Orthohepadnaviridae hepatitis B virus
US-09-239-043D-1667

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 32
US-09-870-089B-13
/ Sequence 13, Application US/09870089B
/ Patent No. 6737062
/ GENERAL INFORMATION:
/ APPLICANT: Charles A. Nicolette
/ TITLE OF INVENTION: THERAPEUTIC COMPOUNDS FOR OVARIAN CANCER
/ FILE REFERENCE: 68126881209900
/ CURRENT APPLICATION NUMBER: US/09/870,089B
/ CURRENT FILING DATE: 2001-05-30
/ NUMBER OF SEQ ID NOS: 14
/ SOFTWARE: FastSEQ for Windows Version 3.0
/ SEQ ID NO 13
/ LENGTH: 9
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ FEATURE:
/ OTHER INFORMATION: ATF4/CREB-2
US-09-870-089B-13

Query Match          91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      5 PHG 7

RESULT 33
US-09-919-048-102
/ Sequence 102, Application US/09919048
/ Patent No. 6787354
/ GENERAL INFORMATION:
/ APPLICANT: O'Brien, Timothy J.
/ APPLICANT: Cannon, Martin J.
/ APPLICANT: Santin, Alessandro
/ TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
/ FILE REFERENCE: D6223CIP/A/D/CIP
/ CURRENT APPLICATION NUMBER: US/09/919,048
/ CURRENT FILING DATE: 2001-07-30
/ PRIOR APPLICATION NUMBER: 09/861,966
/ PRIOR FILING DATE: 2001-05-21
```

; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 102
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-102

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

RESULT 34
US-09-919-048-124
; Sequence 124, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 124
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-124

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

RESULT 35
US-09-919-048-125
; Sequence 125, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 125
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-919-048-125

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 4 PHG 6

RESULT 36
US-09-919-048-140
; Sequence 140, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 140
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 222-230 of the hepsin protein
US-09-919-048-140

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

RESULT 37
US-09-919-048-144
; Sequence 144, Application US/09919048
; Patent No. 6787354
; GENERAL INFORMATION:
; APPLICANT: O'Brien, Timothy J.
; APPLICANT: Cannon, Martin J.
; APPLICANT: Santin, Alessandro
; TITLE OF INVENTION: Methods for the early diagnosis of ovarian cancer
; FILE REFERENCE: D6223CIP/A/D/CIP
; CURRENT APPLICATION NUMBER: US/09/919,048
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 09/861,966
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 188
; SEQ ID NO 144
; LENGTH: 9
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Residues 224-232 of the hepsin protein
US-09-919-048-144

Query Match 91.3%; Score 21; DB 4; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 4 PHG 6

RESULT 38
US-08-049-511-10
; Sequence 10, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H
; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 92008A
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-049-511-10

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 8 PHG 10

RESULT 39
US-08-049-511-11
; Sequence 11, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H

; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
US-08-049-511-11

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 8 PHG 10

RESULT 40
US-08-049-511-13
; Sequence 13, Application US/08049511
; Patent No. 5316775
; GENERAL INFORMATION:
; APPLICANT: WAGLE, SUDHAKAR S
; APPLICANT: STEINBACH, THOMAS
; APPLICANT: LAWYER, CARL H
; APPLICANT: HERMANN, WILLIAM J
; APPLICANT: GAWISH, ALI ABDEL SALAM
; TITLE OF INVENTION: METHOD OF TREATING
; TITLE OF INVENTION: HEPATITIS B INFECTION
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: TILTON, FALLON, LUNGUMUS & CHESTNUT
; STREET: 100 SOUTH WACKER DRIVE, SUITE 960

;
; CITY: CHICAGO
; STATE: ILLINOIS
; COUNTRY: USA
; ZIP: 60606-4002
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/049,511
; FILING DATE: 19930419
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/780,084
; FILING DATE: 15-OCT-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/804,844
; FILING DATE: 04-DEC-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/728,267
; FILING DATE: 11-JUL-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/228,364
; FILING DATE: 04-AUG-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: FENTRESS, SUSAN B
; REGISTRATION NUMBER: 31,327
; REFERENCE/DOCKET NUMBER: 92008A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312/456-8000
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; US-08-049-511-13

Query Match 91.3%; Score 21; DB 1; Length 10;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 PHG 5
|
|
|
Db 8 PHG 10

Search completed: June 15, 2005, 14:24:18
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-7
Perfect score: 23
Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
 - 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
 - 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
 - 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
 - 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	91.3	4	17 US-10-476-861A-31	Sequence 31, Appl
2	21	91.3	4	17 US-10-923-940-25	Sequence 25, Appl
3	21	91.3	5	9 US-09-817-661-32	Sequence 32, Appl
4	21	91.3	5	14 US-10-074-225A-7	Sequence 7, Appl
5	21	91.3	5	14 US-10-074-225A-8	Sequence 8, Appl
6	21	91.3	5	14 US-10-074-225A-9	Sequence 9, Appl
7	21	91.3	5	14 US-10-074-225A-10	Sequence 10, Appl
8	21	91.3	5	14 US-10-074-225A-11	Sequence 11, Appl
9	21	91.3	6	9 US-09-932-161-5	Sequence 5, Appl
10	21	91.3	6	13 US-10-156-820-71	Sequence 71, Appl
11	21	91.3	6	13 US-10-026-237-25	Sequence 25, Appl

12	21	91.3	6	14 US-10-153-312A-1	Sequence 1, Appl
13	21	91.3	6	14 US-10-096-777-5	Sequence 5, Appl
14	21	91.3	6	16 US-10-699-088-232	Sequence 232, App
15	21	91.3	6	16 US-10-699-088-240	Sequence 240, App
16	21	91.3	6	16 US-10-699-088-758	Sequence 758, App
17	21	91.3	6	16 US-10-699-088-765	Sequence 765, App
18	21	91.3	6	16 US-10-699-088-957	Sequence 957, App
19	21	91.3	6	16 US-10-699-088-964	Sequence 964, App
20	21	91.3	6	16 US-10-699-113-46	Sequence 46, Appl
21	21	91.3	6	16 US-10-699-113-53	Sequence 53, Appl
22	21	91.3	6	16 US-10-699-113-232	Sequence 232, App
23	21	91.3	6	16 US-10-699-113-240	Sequence 240, App
24	21	91.3	6	16 US-10-699-113-758	Sequence 758, App
25	21	91.3	6	16 US-10-699-113-765	Sequence 765, App
26	21	91.3	6	17 US-10-699-114-232	Sequence 232, App
27	21	91.3	6	17 US-10-699-114-240	Sequence 240, App
28	21	91.3	6	17 US-10-699-114-758	Sequence 758, App
29	21	91.3	6	17 US-10-699-114-765	Sequence 765, App
30	21	91.3	6	17 US-10-699-114-957	Sequence 957, App
31	21	91.3	6	17 US-10-699-114-964	Sequence 964, App
32	21	91.3	6	17 US-10-806-924-9	Sequence 9, Appl
33	21	91.3	6	17 US-10-806-924-16	Sequence 16, Appl
34	21	91.3	6	17 US-10-806-924-195	Sequence 195, App
35	21	91.3	6	17 US-10-806-924-203	Sequence 203, App
36	21	91.3	6	17 US-10-806-924-721	Sequence 721, App
37	21	91.3	6	17 US-10-806-924-728	Sequence 728, App
38	21	91.3	7	10 US-09-954-385-70	Sequence 70, Appl
39	21	91.3	7	15 US-10-646-807-12	Sequence 12, Appl
40	21	91.3	7	15 US-10-912-512-70	Sequence 70, Appl
41	21	91.3	7	17 US-10-912-512-377	Sequence 377, App
42	21	91.3	7	17 US-10-235-043-70	Sequence 70, Appl
43	21	91.3	7	17 US-10-235-043-377	Sequence 377, App
44	21	91.3	8	9 US-09-734-357-2	Sequence 2, Appl
45	21	91.3			

ALIGNMENTS

RESULT 1
US-10-476-861A-31
; Sequence 31, Application US/10476861A
; Publication No. US20050084902A1
; GENERAL INFORMATION:
; APPLICANT: NOKIHARA , Kiyoshi et al.
; TITLE OF INVENTION: PEPTIDE-IMMOBILIZED SUBSTRATE AND METHOD FOR MEASURING TARGET
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 0760-0325P
; CURRENT APPLICATION NUMBER: US/10/476,861A
; CURRENT FILING DATE: 2003-11-06
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 31
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Peptide fragment of SEQ ID
; OTHER INFORMATION: NO: 7 and example 3
US-10-476-861A-31

Query Match 91.3%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db |||
2 PHG 4

RESULT 2
US-10-923-940-25
; Sequence 25, Application US/10923940

```
; Publication No. US20050113297A1
; APPLICANT: Francois, et al.
; TITLE OF INVENTION: Compositions and Methods for Enhancing Phagocytosis or Phagocyte
; FILE REFERENCE: 2005284-0010
; CURRENT FILING DATE: 2004-08-23
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 25
; LENGTH: 4
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Identified by phage display. Not from any organism
US-10-923-940-25

Query Match          91.3%; Score 21; DB 17; Length 4;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 3
US-09-817-661-32
; Sequence 32, Application US/09817661
; Patent No. US20020076692A1
; GENERAL INFORMATION:
; APPLICANT: Osbourn, Jane
; APPLICANT: Holet, Thor
; TITLE OF INVENTION: Improvements to ribosome display
; FILE REFERENCE: 84633
; CURRENT APPLICATION NUMBER: US/09/817,661
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: US 60/193,802
; PRIOR FILING DATE: 2000-03-31
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 32
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Mutagenized
US-09-817-661-32

Query Match          91.3%; Score 21; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      2 PHG 4

RESULT 4
US-10-074-225A-7
; Sequence 7, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
```

```
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
; NAME/KEY: MISC_FEATURE
; LOCATION: (1)..(2)
; OTHER INFORMATION: Xaa at positions 1 and 2 can be either His or Pro
US-10-074-225A-7

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      3 PHG 5

RESULT 5
US-10-074-225A-8
; Sequence 8, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-8

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      3 PHG 5

RESULT 6
US-10-074-225A-9
; Sequence 9, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
```

```
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-9

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      3 PHG 5

RESULT 7
US-10-074-225A-10
; Sequence 10, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-10

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      3 PHG 5

RESULT 8
US-10-074-225A-11
; Sequence 11, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
```

```
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 11
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-11

Query Match          91.3%; Score 21; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      3 PHG 5

RESULT 9
US-09-932-161-5
; Sequence 5, Application US/09932161
; Patent No. US20020037533A1
; GENERAL INFORMATION:
; APPLICANT: Civealli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Screening and Therapeutic Methods For
; FILE REFERENCE: P-UC 4679
; CURRENT APPLICATION NUMBER: US/09/932,161
; CURRENT FILING DATE: 2001-08-17
; PRIOR APPLICATION NUMBER: US 09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human GPR10 variant
US-09-932-161-5

Query Match          91.3%; Score 21; DB 9; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
Db      1 PHG 3

RESULT 10
US-10-156-820-71
; Sequence 71, Application US/10156820
; Publication No. US20020150558A1
; GENERAL INFORMATION:
; APPLICANT: Boulanger, Pierre
; APPLICANT: Hong, Saw See
; APPLICANT: Karayan, Lucie
; TITLE OF INVENTION: Use of a Polypeptide as Cell Receptor for Adenoviruses
; FILE REFERENCE: 032751-036
; CURRENT APPLICATION NUMBER: US/10/156,820
; CURRENT FILING DATE: 2002-06-30
; PRIOR APPLICATION NUMBER: PCT/FR98/00184
; PRIOR FILING DATE: 1998-01-30
; PRIOR APPLICATION NUMBER: FR 97/01005
; PRIOR FILING DATE: 1997-01-30
; PRIOR APPLICATION NUMBER: FR 97/11166
; PRIOR FILING DATE: 1997-09-09
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: FastSeq for Windows Version 4.0
```

; SEQ ID NO 71
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Phagotope
US-10-156-820-71

Query Match 91.3%; Score 21; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 11
US-10-026-237-25
; Sequence 25, Application US/10026237
; Publication No. US20020173452A1
; GENERAL INFORMATION:
; APPLICANT: Bayer AG
; TITLE OF INVENTION: Cytostatic-Glycoconjugates having specifically cleavable linking
; FILE REFERENCE: Lea 34 491
; CURRENT APPLICATION NUMBER: US/10/026,237
; CURRENT FILING DATE: 2002-10-31
; PRIOR APPLICATION NUMBER: EP 00128402.5
; PRIOR FILING DATE: 2000-12-27
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: Amino Acid sequence not obtained from organism, but synthesized in
US-10-026-237-25

Query Match 91.3%; Score 21; DB 13; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 12
US-10-153-312A-1
; Sequence 1, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-1

Query Match 91.3%; Score 21; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 4 PHG 6

RESULT 13
US-10-096-777-5
; Sequence 5, Application US/10096777
; Publication No. US20030171270A1
; GENERAL INFORMATION:
; APPLICANT: Civelli, Olivier
; APPLICANT: Lin, Steven
; TITLE OF INVENTION: Therapeutic Compositions and Methods
; FILE REFERENCE: P-UC 3534
; CURRENT APPLICATION NUMBER: US/10/096,777
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: US/09/560,915
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: human GPR10 variant
US-10-096-777-5

Query Match 91.3%; Score 21; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 14
US-10-699-088-232
; Sequence 232, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-232

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 15
US-10-699-088-240
; Sequence 240, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-240

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 16
US-10-699-088-758
; Sequence 758, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-758

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 17
US-10-699-088-765
; Sequence 765, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-765

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 3 PHG 5

RESULT 18
US-10-699-088-957
; Sequence 957, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 957
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-957

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 19

```

US-10-699-088-964
; Sequence 964, Application US/10699088
; Publication No. US20040209282A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; TITLE OF INVENTION: METHODS FOR PRODUCING POLYPEPTIDE-TAGGED COLLECTIONS AND CAPTURE
; FILE OF INVENTION: SYSTEMS CONTAINING THE TAGGED POLYPEPTIDES
; FILE REFERENCE: 25885-1754
; CURRENT APPLICATION NUMBER: US/10/699,088
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-088-964

```

```

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 PHG 5
Db 3 PHG 5

```

```

RESULT 20
US-10-699-113-46
; Sequence 46, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-46

```

```

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 PHG 5
Db 2 PHG 4

```

```

RESULT 21
US-10-699-113-53
; Sequence 53, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:

```

```

; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 53
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-53

```

```

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 PHG 5
Db 3 PHG 5

```

```

RESULT 22
US-10-699-113-232
; Sequence 232, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-232

```

```

Query Match          91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 PHG 5
Db 1 PHG 3

```

```

RESULT 23
US-10-699-113-240
; Sequence 240, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof

```

; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-240

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 24

US-10-699-113-758
; Sequence 758, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10
; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-758

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 2 PHG 4

RESULT 25

US-10-699-113-765
; Sequence 765, Application US/10699113
; Publication No. US20040241748A1
; GENERAL INFORMATION:
; APPLICANT: Ault-Riche, Dana
; APPLICANT: Kumble, Krishnanand
; APPLICANT: Schulz, Rainer
; APPLICANT: Schulz, Kenneth
; TITLE OF INVENTION: Self-Assembling Arrays and Uses Thereof
; FILE REFERENCE: 25885-1755
; CURRENT APPLICATION NUMBER: US/10/699,113
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/446,687
; PRIOR FILING DATE: 2003-02-10

; NUMBER OF SEQ ID NOS: 948
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-113-765

Query Match 91.3%; Score 21; DB 16; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 3 PHG 5

RESULT 26

US-10-699-114-232
; Sequence 232, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 232
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-232

Query Match 91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 1 PHG 3

RESULT 27

US-10-699-114-240
; Sequence 240, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018

```
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 240
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-240

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      1 PHG 3

RESULT 28
US-10-699-114-758
; Sequence 758, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 758
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-758

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      2 PHG 4

RESULT 29
US-10-699-114-765
; Sequence 765, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE REFERENCE: 25885-1759
```

```
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 765
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-765

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      3 PHG 5

RESULT 30
US-10-699-114-957
; Sequence 957, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 957
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-957

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
DB      2 PHG 4

RESULT 31
US-10-699-114-964
; Sequence 964, Application US/10699114
; Publication No. US20050042623A1
; GENERAL INFORMATION:
; APPLICANT: Dana Ault-Riche
; APPLICANT: Bruce Atkinson
; APPLICANT: Krishnanand Kumble
; APPLICANT: Lynne Jersaitis
; APPLICANT: Gizette Sperinde
```

```
; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-964

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 32
US-10-806-924-9
; Sequence 9, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 9
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-9

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 33
US-10-806-924-16
; Sequence 16, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0

; TITLE OF INVENTION: SYSTEMS FOR CAPTURE AND ANALYSIS OF BIOLOGICAL PARTICLES AND METH
; TITLE OF INVENTION: USING THE SYSTEMS
; FILE REFERENCE: 25885-1759
; CURRENT APPLICATION NUMBER: US/10/699,114
; CURRENT FILING DATE: 2003-10-30
; PRIOR APPLICATION NUMBER: 60/423,018
; PRIOR FILING DATE: 2002-10-30
; PRIOR APPLICATION NUMBER: 60/422,923
; PRIOR FILING DATE: 2002-10-30
; NUMBER OF SEQ ID NOS: 1094
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 964
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-699-114-964

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 34
US-10-806-924-195
; Sequence 195, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-195

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 35
US-10-806-924-203
; Sequence 203, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-203

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3
```

```
; SEQ ID NO 16
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-16

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 34
US-10-806-924-195
; Sequence 195, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 195
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-195

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 35
US-10-806-924-203
; Sequence 203, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 203
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-203

Query Match          91.3%; Score 21; DB 17; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      1 PHG 3
```

```
QY      3 PHG 5
      |||
Db      1 PHG 3

RESULT 36
US-10-806-924-721
; Sequence 721, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 721
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-721

      Query Match      91.3%; Score 21; DB 17; Length 6;
      Best Local Similarity 100.0%; Pred. No. 1.5e+06;
      Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      2 PHG 4

RESULT 37
US-10-806-924-728
; Sequence 728, Application US/10806924
; Publication No. US20050095648A1
; GENERAL INFORMATION:
; APPLICANT: Geysen, H. Mario
; APPLICANT: Ault-Riche, Dana
; TITLE OF INVENTION: Method for designing linear epitopes and algorithm therefor
; TITLE OF INVENTION: and polypeptide epitopes
; FILE REFERENCE: 25885-1760
; CURRENT APPLICATION NUMBER: US/10/806,924
; CURRENT FILING DATE: 2004-03-22
; NUMBER OF SEQ ID NOS: 911
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 728
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic peptide
US-10-806-924-728

      Query Match      91.3%; Score 21; DB 17; Length 6;
      Best Local Similarity 100.0%; Pred. No. 1.5e+06;
      Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      3 PHG 5

RESULT 38
US-09-954-385-70
; Sequence 70, Application US/09954385
; Publication No. US20030100467A1
; GENERAL INFORMATION:
; APPLICANT: Aehle, Wolfgang
; APPLICANT: Baldwin, Toby L.
; APPLICANT: Van Gastel, Franciscus J.C.
; APPLICANT: Janssen, Giselle G.
; APPLICANT: Murray, Christopher J.
; APPLICANT: Wang, Huaming
; APPLICANT: Winetzk, Deborah S.
; TITLE OF INVENTION: Binding Phenol Oxidizing Enzyme-peptide
; TITLE OF INVENTION: Complexes
; FILE REFERENCE: GC690
; CURRENT APPLICATION NUMBER: US/09/954,385
; CURRENT FILING DATE: 2001-09-12
; NUMBER OF SEQ ID NOS: 433
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 70
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: binding peptide
US-09-954-385-70

      Query Match      91.3%; Score 21; DB 10; Length 7;
      Best Local Similarity 100.0%; Pred. No. 1.5e+06;
      Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 PHG 5
      |||
Db      5 PHG 7

RESULT 40
US-10-646-807-12
; Sequence 12, Application US/10646807
; Publication No. US20040077005A1
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
```

;
; CAMERON, Beatrice
; CROUZET, Joel
; FAMECHON, Alain
; FERRERO, Lucia
; TITLE OF INVENTION: Novel Topoisomerase IV, Corresponding
; Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/646,807
; FILING DATE: 25-Aug-2003
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
; US-10-646-807-12

Query Match 91.3%; Score 21; DB 15; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 3 PHG 5

Search completed: June 15, 2005, 14:50:17
Job time : 80.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-7

Perfect score: 23

Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	21	91.3	18	2	I78841	thrombopoietin rec
2	21	91.3	23	2	A44626	dihydropyrimidine
3	21	91.3	27	2	A27108	larval-specific li
4	21	91.3	32	2	S11610	ribosomal protein
5	21	91.3	34	2	C31514	hemopexin - chicke
6	21	91.3	36	2	E41080	rbpL protein - Rho
7	21	91.3	38	2	G75398	conserved hypotet
8	21	91.3	40	2	S12207	hypothetical prote
9	21	91.3	41	2	A71278	hypothetical prote
10	21	91.3	42	2	H18887	hypothetical prote
11	21	91.3	44	2	H81370	SOS ribosomal prot
12	21	91.3	46	2	A98802	hypothetical prote
13	21	91.3	46	2	E97985	hypothetical prote
14	21	91.3	48	2	T50885	hypothetical solub
15	21	91.3	52	2	H81124	hypothetical prote
16	21	91.3	54	2	S66330	protein kinase AK1
17	21	91.3	56	2	B86823	hypothetical prote
18	21	91.3	57	2	S57791	probable beta-gluc
19	21	91.3	57	2	T49863	hypothetical prote
20	21	91.3	57	2	A33811	tracheobronchial m
21	21	91.3	57	2	C97837	(p)ppGpp 3-pyropho
22	21	91.3	58	2	A45765	cre protein - Stre
23	21	91.3	59	2	T03588	pyrophosphate-fruc
24	21	91.3	59	2	A86485	protein F28J9.16 l
25	21	91.3	59	2	D82459	hypothetical prote
26	21	91.3	60	2	A84213	hypothetical prote
27	21	91.3	60	2	G82472	hypothetical prote
28	21	91.3	61	1	ERAD64	early E3 6.4K prot
29	21	91.3	64	2	E86800	prophage pi3 prote

ALIGNMENTS

RESULT 1

I78841
thrombopoietin receptor - mouse (fragment)

C:Species: Mus sp. (mouse)

C>Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999

C:Accession: I78841

R:Alexander, W.S.; Dunn, A.R.

Oncogene 10, 795-803, 1995

A>Title: Structure and transcription of the genomic locus encoding murine c-Mpl, a receptor for thrombopoietin; MUID:95166571; PMID:7862460

A:Reference number: I58350

A:Accession: I78841

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-18 <RES>

A:Cross-references: GB:S76842; NID:g912990; PIDN:AAB33462.1; PID:g912991

C:Genetics:

A:Gene: c-mpl

Query Match 91.3%; Score 21; DB 2; Length 18;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

DB 2 PHG 4

RESULT 2

A4626

dihydropyrimidine dehydrogenase (NADP) (EC 1.3.1.2) - bovine (fragment)

N:Alternate names: dihydrothymine dehydrogenase; dihydrouracil dehydrogenase (NADP+)

C:Species: Bos primigenius taurus (cattle)

C>Date: 02-Aug-1994 #sequence_revision 23-Mar-1995 #text_change 03-Jun-2002

C:Accession: A44626; A41066

R:Porter, D.J.T.; Chestnut, W.G.; Merrill, B.M.; Spector, T.

J. Biol. Chem. 267, 5236-5242, 1992

A>Title: Mechanism-based inactivation of dihydropyrimidine dehydrogenase by 5-ethynyluracil

A:Reference number: A44626; MUID:92184771; PMID:1544906

A:Accession: A44626

A>Status: preliminary

A:Molecule type: protein

A:Residues: 1-23 <POR>

R:Porter, D.J.T.; Chestnut, W.G.; Taylor, L.C.E.; Merrill, B.M.; Spector, T.

J. Biol. Chem. 266, 19988-19994, 1991

A>Title: Inactivation of dihydropyrimidine dehydrogenase by 5-iodouracil.

A:Reference number: A41066; MUID:92041818; PMID:1939061

A:Accession: A41066

A:Molecule type: protein

A:Residues: 13-23 <PO2>

C:Superfamily: unassigned ferredoxin 2[4Fe-4S]-related proteins; ferredoxin 2[4Fe-4S] hom

C:Keywords: NADP; oxidoreductase

F16/Binding site: substrate (Cys) #status experimental

Query Match 91.3%; Score 21; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 17 PHG 19

RESULT 3
A27108
larval-specific lipoprotein - honeybee (fragment)
C:Species: Apis mellifera (honeybee)
C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 09-Jul-2004
C:Accession: A27108
R:Shipman, B.A.; Ryan, R.O.; Schmidt, J.O.; Law, J.H.
Biochemistry 26, 1885-1889, 1987
A:Title: Purification and properties of a very high density lipoprotein from the hemolymph of the honeybee, Apis mellifera (L.)
A:Reference number: A27108; MUID:87242376; PMID:3109474
A:Accession: A27108
A:Molecule type: protein
A:Residues: 1-27 <SH1>
A:Cross-references: UNIPROT:P09355
C:Keywords: lipid binding; lipoprotein

Query Match 91.3%; Score 21; DB 2; Length 27;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 5 PHG 7

RESULT 4
S11610
ribosomal protein S4.er [validated] - Halobacterium salinarum (fragment)
N:Alternate names: ribosomal protein HS5
C:Species: Halobacterium salinarum
C:Date: 31-Dec-1990 #sequence_revision 31-Dec-1990 #text_change 09-Jul-2004
C:Accession: S11610
R:Yaguchi, M.; Visentin, L.P.; Zuker, M.; Matheson, A.T.; Roy, C.; Strom, A.R.
Zbl. Bakt. Hyg. I. Abt. Orig. C 3, 200-208, 1982
A:Title: Amino-terminal sequences of ribosomal proteins from the 30S subunit of archaebacteria
A:Reference number: S11609
A:Accession: S11610
A:Molecule type: protein
A:Residues: 1-32 <YAG>
A:Cross-references: UNIPROT:Q7M553
A:Note: the protein is designated as ribosomal protein HS5
A:Note: the source is designated as Halobacterium cutirubrum
C:Superfamily: rat ribosomal protein S4
C:Keywords: protein biosynthesis; ribosome

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 30 PHG 32

RESULT 5
C31514
hemopexin - chicken (fragment)
C:Species: Gallus gallus (chicken)
C:Date: 28-Feb-1990 #sequence_revision 28-Feb-1990 #text_change 09-Jul-2004
C:Accession: C31514
R:Wellner, D.; Cheng, K.C.; Muller-Eberhard, U.
Biochem. Biophys. Res. Commun. 155, 622-625, 1988
A:Title: N-terminal amino acid sequences of the hemopexins from chicken, rat and rabbit.

A:Reference number: A90148; MUID:88339942; PMID:3421961
A:Accession: C31514
A:Molecule type: protein
A:Residues: 1-34 <WE3>
A:Cross-references: UNIPROT:P20057

Query Match 91.3%; Score 21; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 16 PHG 18

RESULT 6
E41080
rpbL protein - Rhodobacter sphaeroides (fragment)
C:Species: Rhodobacter sphaeroides
C:Date: 12-Jun-1992 #sequence_revision 12-Jun-1992 #text_change 09-Jul-2004
C:Accession: E41080
R:Chen, J.H.; Gibson, J.L.; McCue, L.A.; Tabita, F.R.
J. Biol. Chem. 266, 20447-20452, 1991
A:Title: Identification, expression, and deduced primary structure of transketolase and transketolase-related proteins from Rhodobacter sphaeroides
A:Reference number: A41080; MUID:92041881; PMID:1939098
A:Accession: E41080
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-36 <CHE>
A:Cross-references: UNIPROT:P29278; GB:M68914; NID:gl51988; PIDN:AAA26158.1; PID:g551958

Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 25 PHG 27

RESULT 7
G75398
conserved hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: G75398
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; F.
S.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896; PMID:10567266
A:Accession: G75398
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-38 <WH1>
A:Cross-references: UNIPROT:O9RUH4; GB:AE001986; GB:AE000513; NID:g6459162; PIDN:AAF1097f
A:Experimental source: strain R1
C:Genetics:
A:Gene: DR1412
A:Map position: 1

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 31 PHG 33

RESULT 8
S12207

Y, P.; Sun, P.M.; Winkler, M.E.
J. Bacteriol. 183, 5709-5717, 2001
A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
A:Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
A:Reference number: A97872; MUID:21429245; PMID:11544234
A:Accession: E97985
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-46 <KUR>
A:Cross-references: UNIPROT:Q8CYU7; GB:AE007317; PIDN:AAK99713.1; PID:g15458516; GSPDB:G
C:Genetics:
A:Gene: spr0909

Query Match 91.3%; Score 21; DB 2; Length 46;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 23 PHG 25

RESULT 14
T50885
hypothetical soluble protein [imported] - Rubrivivax gelatinosus
C:Species: Rubrivivax gelatinosus
C:Date: 21-Jul-2000 #sequence_revision 21-Jul-2000 #text_change 09-Jul-2004
C:Accession: T50885
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shimada, K.
submitted to the EMBL data library, November 1999
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photosyn
A:Reference number: Z25270
A:Accession: T50885
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-48 <NAG>
A:Cross-references: UNIPROT:Q9JPB8; EMBL:AB034704; PIDN:BAA94038.1
A:Experimental source: strain IL14
C:Genetics:
A:Note: ORF48

Query Match 91.3%; Score 21; DB 2; Length 48;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 46 PHG 48

RESULT 15
H81124
hypothetical protein NMB1079 [imported] - Neisseria meningitidis (strain MC58 serogroup
C:Species: Neisseria meningitidis
C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: H81124
R:Tetzelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ve
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: H81124
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-52 <TET>
A:Cross-references: UNIPROT:Q9JZF4; GB:AE002458; PIDN:AAF4147
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1079

Query Match 91.3%; Score 21; DB 2; Length 52;

Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 13 PHG 15

RESULT 16
S66330
protein kinase AKL7 (EC 2.7.1.1-) - Arabidopsis thaliana (fragment)
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 16-Aug-2004
C:Accession: S66330; S58271
R:Thummler, F.; Kirchner, M.; Teuber, R.; Dittrich, P.
Plant Mol. Biol. 29, 551-565, 1995
A:Title: Differential accumulation of the transcripts of 22 novel protein kinase genes in
A:Reference number: S66314; MUID:96123233; PMID:8534852
A:Accession: S66330
A:Molecule type: DNA
A:Residues: 1-54 <THU>
A:Cross-references: UNIPROT:Q38982; EMBL:X86962; NID:g928899; PIDN:CAA60525.1; PID:g9289
C:Genetics:
A:Gene: AKL7
C:Superfamily: protein kinase homology
C:Keywords: ATP; phosphotransferase; protein kinase
F,1-54/Domain: protein kinase homology (fragment) <KIN>

Query Match 91.3%; Score 21; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 49 PHG 51

RESULT 17
B86823
hypothetical protein yqcG [imported] - Lactococcus lactis subsp. lactis (strain IL1403)
C:Species: Lactococcus lactis subsp. lactis
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004
C:Accession: B86823
R:Boilotin, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissenbach, J.; Ehrlic
Genome Res. 11, 731-753, 2001
A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis ser
A:Reference number: AB6625; MUID:21235186; PMID:11337471
A:Accession: B86823
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-56 <STO>
A:Cross-references: UNIPROT:Q9CF95; GB:AE005176; PID:g12724591; PIDN:AAK05684.1; GSPDB:G
A:Experimental source: strain IL1403
C:Genetics:
A:Gene: yqcG

Query Match 91.3%; Score 21; DB 2; Length 56;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
DB 36 PHG 38

RESULT 18
S57791
probable beta-glucosidase - black mustard (fragment)
C:Species: Brassica nigra (black mustard)
C:Date: 27-Oct-1995 #sequence_revision 10-May-1996 #text_change 09-Jul-2004
C:Accession: S57791
R:Malboodi, M.A.; Lefebvre, D.D.
Plant Mol. Biol. 28, 859-870, 1995

A;Title: Isolation of cDNA clones of genes with altered expression levels in phosphate-de
A;Reference number: S57791; MUID:95367647; PMID:7640358
A;Accession: S57791
A;Status: nucleic acid sequence not shown; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-57 <MAL>
A;Cross-references: UNIPROT:O24434
C;Superfamily: Agrobacterium beta-glucosidase

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 30 PHG 32

RESULT 19
T49863
hypothetical protein B24P11.190 [imported] - Neurospora crassa
C;Species: Neurospora crassa
C;Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
C;Accession: T49863
R;Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,
submitted to the Protein Sequence Database, May 2000
A;Reference number: Z25022
A;Accession: T49863
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <SCH>
A;Cross-references: EMBL:AL356833; GSPDB:GN00116; NCSP:B24P11.190
A;Experimental source: BAC clone B24P11; strain OR74A
C;Genetics:
A;Gene: NCSP:B24P11.190
A;Map position: 6
A;Introns: 9/1

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 35 PHG 37

RESULT 20
A33811
tracheobronchial mucin - human (fragments)
C;Species: Homo sapiens (man)
C;Date: 09-Mar-1990 #sequence_revision 09-Mar-1990 #text_change 09-Jul-2004
C;Accession: A33811
R;Rose, M.C.; Kaufman, B.; Martin, B.M.
J. Biol. Chem. 264, 8193-8199, 1989

A;Title: Proteolytic fragmentation and peptide mapping of human carboxyamidomethylated b
A;Reference number: A33811; MUID:89255255; PMID:2656675
A;Accession: A33811
A;Status: preliminary
A;Molecule type: protein
A;Residues: 1-57 <ROS>
A;Cross-references: UNIPROT:Q7M4S5

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 11 PHG 13

RESULT 21

C97837
(p)ppGpp 3-pyrophosphohydrolase homolog RC1099 [imported] - Rickettsia conorii (strain Me
C;Species: Rickettsia conorii
C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C;Accession: C97837
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro
Science 293, 2093-2098, 2001
A;Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A;Reference number: A97700; MUID:21442074; PMID:11557893
A;Accession: C97837
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-57 <KUR>
A;Cross-references: UNIPROT:Q92GM4; GB:AE006914; PIDN:AAL03637.1; PID:g15620223; GSPDB:G
C;Genetics:
A;Gene: RC1099

Query Match 91.3%; Score 21; DB 2; Length 57;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 39 PHG 41

RESULT 22
A45765
cre protein - Streptomyces hygroscopicus
C;Species: Streptomyces hygroscopicus
C;Date: 14-May-1993 #sequence_revision 14-May-1993 #text_change 09-Jul-2004
C;Accession: A45765
R;Ogura, M.; Tanaka, T.; Seto, H.; Otake, N.
J. Antibiot. 43, 873-882, 1990
A;Title: Molecular cloning and characterization of the gene conferring curromycin resist
A;Reference number: A45765; MUID:90354306; PMID:2117601
A;Accession: A45765
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-58 <OGU>
A;Cross-references: UNIPROT:P16961; GB:M28599; NID:gi53217; PIDN:AAA26721.1; PID:g153218

Query Match 91.3%; Score 21; DB 2; Length 58;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 14 PHG 16

RESULT 23
T03588
pyrophosphate-fructose-6-phosphate 1-phosphotransferase homolog - rice (fragment)
C;Species: Oryza sativa (rice)
C;Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C;Accession: T03588
R;Umeda, M.; Hara, C.; Matsubayashi, Y.; Li, H.; Lui, Q.; Tadokoro, F.; Aotsuka, S.; Uchi
Plant Mol. Biol. 25, 469-478, 1994
A;Title: Expressed sequence tags from cultured cells of rice (Oryza sativa L.) under stre
A;Reference number: Z14962; MUID:94325471; PMID:8049371
A;Accession: T03588
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-59 <UME>
A;Cross-references: UNIPROT:Q42988; EMBL:D17765; NID:g455496; PIDN:BAA04611.1; PID:g87191
A;Experimental source: callus
C;Genetics:
A;Note: AK99
C;Superfamily: pyrophosphate-dependent phosphofructokinase, PfpB type; 6-phosphofructokir

Query Match 91.3%; Score 21; DB 2; Length 59;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;

Matches	3;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	3 PHG 5								
Db	26 PHG 28								
RESULT 24									
A:Accession: A86485	protein F28J9.16 [imported] - Arabidopsis thaliana								
C:Species: Arabidopsis thaliana (mouse-ear cress)									
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001	#text_change 09-Jul-2004								
A:Reference number: A86485									
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.									
Nature 408, 815-820, 2000									
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.									
A:Authors: Salzberg, S.L.; Schwaartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.									
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.									
A:Reference number: A86141; PMID:21016719; PMID:11130712									
A:Accession: A86485									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-59 <STO>									
A:Cross-references: UNIPROT:Q9S9R3; GB:AE005172; NID:g6272386; PIDN:AAF06092.1; GSPDB:GN									
C:Genetics:									
A:Gene: F28J9.16									
A:Map position: 1									
Query Match	91.3%;	Score 21;	DB 2;	Length 59;					
Best Local Similarity	100.0%;	Pred. No. 4.1e+02;							
Matches	3;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	3 PHG 5								
Db	12 PHG 14								
RESULT 25									
D82459	hypothetical protein VCA0435 [imported] - Vibrio cholerae (strain N16961 serogroup O1)								
C:Species: Vibrio cholerae									
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000	#text_change 09-Jul-2004								
A:Accession: D82459									
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, H									
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.									
Nature 406, 477-483, 2000									
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.									
A:Reference number: A82035; PMID:20406833; PMID:10952301									
A:Accession: D82459									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-59 <HEI>									
A:Cross-references: UNIPROT:Q9KMC7; GB:AE004377; GB:AE003853; NID:g9657831; PIDN:AAF9634									
A:Experimental source: serogroup O1; strain N16961; biotype El Tor									
C:Genetics:									
A:Gene: VCA0435									
A:Map position: 2									
Query Match	91.3%;	Score 21;	DB 2;	Length 59;					
Best Local Similarity	100.0%;	Pred. No. 4.1e+02;							
Matches	3;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	3 PHG 5								
Db	5 PHG 7								

RESULT 26									
A84213	hypothetical protein Vng0548c [imported] - Halobacterium sp. NRC-1								
C:Species: Halobacterium sp. NRC-1									
C>Date: 02-Feb-2001 #sequence_revision 02-Feb-2001	#text_change 09-Jul-2004								
A:Accession: A84213									
R:Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Lasky, S.; Leithauser, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor Jung, K.H.; Alam, M.; Freitas, T.									
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000									
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li									
A>Title: Genome sequence of Halobacterium species NRC-1.									
A:Reference number: A84160; PMID:20504483; PMID:11016950									
A:Accession: A84213									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-60 <STO>									
A:Cross-references: UNIPROT:Q9HRT9; GB:AE004437; NID:g10580147; PIDN:AAG19069.1; GSPDB:GN									
C:Genetics:									
A:Gene: VNG0548C									
Query Match	91.3%;	Score 21;	DB 2;	Length 60;					
Best Local Similarity	100.0%;	Pred. No. 4.2e+02;							
Matches	3;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	3 PHG 5								
Db	46 PHG 48								
RESULT 27									
G82472	hypothetical protein VCA0336 VCA0353 VCA0297 [imported] - Vibrio cholerae (strain N16961								
C:Species: Vibrio cholerae									
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000	#text_change 09-Jul-2004								
A:Accession: G82472; A82471; A82477									
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.; chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P. J.									
l, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.									
Nature 406, 477-483, 2000									
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.									
A:Reference number: A82035; PMID:20406833; PMID:10952301									
A:Accession: G82472									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-60 <HEI>									
A:Cross-references: UNIPROT:O31021; GB:AE004372; GB:AE003853; NID:g9657741; PIDN:AAF9624									
A:Experimental source: serogroup O1; strain N16961; biotype El Tor									
A:Accession: A82471									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-60 <HE2>									
A:Cross-references: GB:AE004372; GB:AE003853; NID:g9657741; PIDN:AAF96261.1; GSPDB:GN001									
A:Experimental source: serogroup O1; strain N16961; biotype El Tor									
A:Accession: A82477									
A>Status: preliminary									
A:Molecule type: DNA									
A:Residues: 1-60 <HE3>									
A:Cross-references: GB:AE004369; GB:AE003853; NID:g9657685; PIDN:AAF96205.1; GSPDB:GN001									
A:Experimental source: serogroup O1; strain N16961; biotype El Tor									
C:Genetics: <GEN1>									
A:Gene: VCA0336									
A:Map position: 2									
C:Genetics: <GEN2>									
A:Gene: VCA0353									
A:Map position: 2									
C:Genetics: <GEN3>									
A:Gene: VCA0297									
A:Map position: 2									
Query Match	91.3%;	Score 21;	DB 2;	Length 60;					
Best Local Similarity	100.0%;	Pred. No. 4.2e+02;							
Matches	3;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;

QY 3 PHG 5
|||
Db 6 PHG 8

RESULT 28

ERAD64

early E3 6.4K protein - human adenovirus 35

C:Species: Mastadenovirus h35 (human adenovirus 35)

A:Note: Host Homo sapiens (man)

C:Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004

C:Accession: B31162

R:Flomenberg, P.R.; Chen, M.; Horwitz, M.S.

J. Virol. 62, 4431-4437, 1988

A:Title: Sequence and genetic organization of adenovirus type 35 early region 3.

A:Reference number: A93039; MUID:89012230; PMID:3172347

A:Accession: B31162

A:Molecule type: DNA

A:Residues: 1-61 <FLO>

A:Cross-references: UNIPROT:P17591; GB:M23195; NID:G516584; PIDN:AAA42436.1; PID:G516586

C:Superfamily: adenovirus early E3 6.4K protein

C:Keywords: early protein

Query Match 91.3%; Score 21; DB 1; Length 61;

Best Local Similarity 100.0%; Pred. No. 4.3e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 22 PHG 24

RESULT 29

E86800

prophage pi3 protein 32 [imported] - Lactococcus lactis subsp. lactis (strain IL1403)

C:Species: Lactococcus lactis subsp. lactis

C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 09-Jul-2004

C:Accession: E86800

R:Boletín, A.; Wincker, P.; Mauger, S.; Jaillon, O.; Malarne, K.; Weissbach, J.; Ehrlich

Genome Res. 11, 731-753, 2001

A:Title: The complete genome sequence of the lactic acid bacterium Lactococcus lactis s

A:Reference number: A86625; MUID:21235186; PMID:11337471

A:Accession: E86800

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-64 <STO>

A:Cross-references: UNIPROT:Q9CFR0; GB:AE005176; PID:G12724393; PIDN:AAK05503.1; GSPDB:C

A:Experimental source: strain IL1403

C:Genetics:

A:Gene: pi332

Query Match

Best Local Similarity 91.3%; Score 21; DB 2; Length 64;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 34 PHG 36

RESULT 30

PQ0277

hypothetical protein 1 - cercopithecine herpesvirus 1 (fragment)

C:Species: cercopithecine herpesvirus 1

C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 09-Jul-2004

C:Accession: PQ0277

R:Killen, A.M.; Harrington, L.; Wall, L.V.M.; Kelly, D.C.

J. Gen. Virol. 73, 195-199, 1992

A:Title: Nucleotide sequence analysis of a homologue of herpes simplex virus type 1 gene

A:Reference number: JQ1406; MUID:92113572; PMID:1309859

A:Accession: PQ0277

A:Molecule type: DNA

N:Alternate names: hypothetical protein ipa-20r

A:Residues: 1-64 <KIL>

A:Cross-references: UNIPROT:P30816; GB:S75996; NID:G243000; PIDN:AAE21001.1; PID:G243001

C:Comment: Simian herpes B virus is an alpha herpesvirus which causes a mild, recurring

Query Match 91.3%; Score 21; DB 2; Length 64;

Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 62 PHG 64

RESULT 31

S10211

hypothetical protein, 7.3K - human adenovirus 41

C:Species: Mastadenovirus h41 (human adenovirus 41)

C:Date: 12-Feb-1993 #sequence_revision 12-Feb-1993 #text_change 09-Jul-2004

C:Accession: S10211

R:Slomenda, S.B.; Pieniazek, N.J.; Velarde Jr., J.; Pieniazek, D.; Luftig, R.B.

Nucleic Acids Res. 18, 3089, 1990

A:Title: Nucleotide sequence of the region coding for 100K and 33K proteins of human ente

A:Reference number: S10206; MUID:90272433; PMID:2349115

A:Accession: S10211

A>Status: translation not shown

A:Molecule type: DNA

A:Residues: 1-65 <SLE>

A:Cross-references: UNIPROT:P23691; EMBL:X52532; NID:G58438; PIDN:CAA36764.1; PID:G58444

Query Match

Best Local Similarity 91.3%; Score 21; DB 2; Length 65;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 24 PHG 26

RESULT 32

G75258

conserved hypothetical protein - Deinococcus radiodurans (strain R1)

C:Species: Deinococcus radiodurans

C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004

C:Accession: G75258

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, B.K.; Peterson, J.D.; Dodson, R.J.; F

S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999

A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.

A:Reference number: AY5250; MUID:20036896; PMID:10567266

A:Accession: G75258

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-65 <WHI>

A:Cross-references: UNIPROT:Q9RRD1; GB:AE002085; GB:AE000513; NID:G6460383; PIDN:AAF12093

A:Experimental source: strain R1

C:Genetics:

A:Gene: DR2561

A:Map position: 1

Query Match

Best Local Similarity 91.3%; Score 21; DB 2; Length 65;

Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 43 PHG 45

RESULT 33

S39675

ywbE protein - Bacillus subtilis

N:Alternate names: hypothetical protein ipa-20r

C:Species: *Bacillus subtilis*
 C:Date: 07-Oct-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004
 C:Accession: S39675; C70051
 R:Glaser, P.; Kunst, F.; Arnaud, M.; Coudart, M.P.; Gonzales, W.; Hullo, M.F.; Ionescu, A.; Rapoport, G.; Danchin, A.
 Mol. Microbiol. 10, 371-384, 1993
 A:Title: *Bacillus subtilis* genome project: cloning and sequencing of the 97 kb region flanking the *gluA* gene
 A:Reference number: S39655; MUID:95020537; PMID:7934828
 A:Accession: S39675
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-65 <GLA>
 A:Cross-references: UNIPROT:P39588; EMBL:X73124; NID:G413923; PIDN:CAA51576.1; PID:G413923
 R>Note: The nucleotide sequence was submitted to the EMBL Data Library, June 1993
 R:Kunst, F.; Ogasawara, N.; Moser, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bortner, C.; Bron, S.; Brouillet, S.; Bruschini, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chikara, A.; Ehrlich, S.D.; Emmerich, P.T.; Entian, K.D.; Ewing, J.; Fabis, C.; Ferrarini, E.; Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maguda, S.; Maueel, Y., M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Fortetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
 A:Reference number: A69580; MUID:98044033; PMID:9384377
 A:Accession: C70051
 A>Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-65 <KUN>
 A:Cross-references: GB:Z99123; GB:AL009126; NID:G2636240; PIDN:CAB15861.1; PID:el186334;
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: ywB

Query Match 91.3%; Score 21; DB 2; Length 65;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 44 PHG 46

RESULT 34
 B31048
 N:Alternative names: phosphotransferase system enzyme II (EC 2.7.1.69), mannitol-specific, factor III - Staphylococcus carnosus
 C:Species: *Staphylococcus carnosus*
 C:Date: 15-Dec-1988 #sequence_revision 15-Dec-1988 #text_change 09-Jul-2004
 C:Accession: B31048
 R:Reiche, B.; Frank, R.; Deutscher, J.; Meyer, N.; Hengstenberg, W.
 Biochemistry 27, 6512-6516, 1988
 A:Title: *Staphylococcus carnosus* phosphoenolpyruvate-dependent phosphotransferase system: purification and homology with the enzyme II (mtII) of *Escherichia coli*.
 A:Reference number: A31048; MUID:89118247; PMID:3064811
 A:Accession: B31048
 A>Status: preliminary
 A:Molecule type: protein
 A:Residues: 1-67 <REI>
 A:Cross-references: UNIPROT:P17876
 C:Superfamily: phosphotransferase system mannitol-specific enzyme II factor III; phosphotransferase
 F:35-67/Domain: phosphotransferase system mannitol-specific enzyme II factor III homolog

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 63 PHG 65

RESULT 35
 T25844
 A:Title: hypothetical protein M03F4.1 - *Caenorhabditis elegans*
 C:Species: *Caenorhabditis elegans*
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
 C:Accession: T25844
 R:Fulton, L.
 submitted to the EMBL Data Library, July 1996
 A:Description: The sequence of *C. elegans* cosmid M03F4.
 A:Reference number: Z20097
 A:Accession: T25844
 A>Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-67 <FUL>
 A:Cross-references: EMBL:U64601; PIDN:AAB04580.1; GSPDB:GN000028; CESP:M03F4.1
 A:Experimental source: strain Bristol N2; clone M03F4
 C:Genetics:
 A:Gene: CESP:M03F4.1
 A:Map position: X

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 28 PHG 30

RESULT 36
 AC1037
 A:Title: probable phage tail protein STY4622 [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A>Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AC1037
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, T.; Connor, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serovar Typhi
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AC1037
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-67 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD06744.1; PID:gl6505395; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY4622

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 ||||
 Db 44 PHG 46

RESULT 37
 AG0927
 A:Title: probable phage tail protein [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
 A>Note: this species has also been called *Salmonella typhi*
 C:Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
 C:Accession: AG0927
 R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,

th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, S.; Moule, S.; O'Gaora, P.
 Nature 413, 848-852, 2001
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov
 A:Reference number: AB0502; MUID:21534947; PMID:11677608
 A:Accession: AG0927
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-67 <PAR>
 A:Cross-references: GB:AL513382; PIDN:CAD09441.1; PID:gl6504558; GSPDB:GN00176
 C:Genetics:
 A:Gene: STY3680

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db 44 PHG 46

RESULT 38
 B55855
 tail protein - phase P2
 C:Species: phase P2
 C:Date: 15-Oct-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
 C:Accession: B55855
 R:Ziermann, R.; Bartlett, B.; Calendar, R.; Christie, G.E.
 J. Bacteriol. 176, 4974-4984, 1994
 A:Title: Functions involved in bacteriophage P2-induced host cell lysis and identificati
 A:Reference number: A55855; MUID:94327465; PMID:8051010
 A:Accession: B55855
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-67 <ZIE>
 A:Cross-references: UNIPROT:P51772; GB:L29304; NID:g508561; PID:g508562
 C:Genetics:
 A:Gene: X

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db 44 PHG 46

RESULT 39
 T17586
 hypothetical protein a96R - Chlorella virus PBCV-1
 C:Species: Chlorella virus PBCV-1
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
 C:Accession: T17586
 R:Graves, M.V.; Van Etten, J.L.
 submitted to the EMBL Data Library, May 1999
 A:Reference number: Z18806
 A:Accession: T17586
 A:Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-67 <GRA>
 A:Cross-references: UNIPROT:Q84417; EMBL:U42580; NID:g4028986; PIDN:AAC96464.1
 A:Experimental source: specific host Chlorella strain NC64A
 C:Genetics:
 A:Gene: a96R

Query Match 91.3%; Score 21; DB 2; Length 67;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5

Db 33 PHG 35

RESULT 40

G95330
 hypothetical protein Sma1017 [imported] - Sinorhizobium meliloti (strain 1021) magaplaasm
 C:Species: Sinorhizobium meliloti
 C:Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C:Accession: G95330
 R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowe
 ; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
 Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A:Reference number: A95262; MUID:21396509; PMID:11481432
 A:Accession: G95330
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-68 <KUR>
 A:Cross-references: UNIPROT:Q922D7; GB:AE006469; PIDN:AAK65209.1; PID:gl4523656; GSPDB:G
 R:Experimental source: strain 1021, megaplasmaid pSymA
 R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
 pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federgpiel, N.A.; Fisher, R.F.;
 L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
 hebaull, P.; Vandenbol, M.; Vorholter, P.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A:Reference number: A96039; MUID:21368234; PMID:11474104
 C:Genetics:
 A:Gene: Sma1017
 A:Genome: plasmid

Query Match 91.3%; Score 21; DB 2; Length 68;
 Best Local Similarity 100.0%; Pred. No. 4.8e+02;
 Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
 Db 48 PHG 50

Search completed: June 15, 2005, 14:22:43
 Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-7
Perfect score: 23
Sequence: 1 XXPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	21	91.3	13	Q80XX0	Q80xx0 cricetus
2	21	91.3	16	O21922	O21922 streptococ
3	21	91.3	18	Q6LD80	Q6ld80 mus sp. thr
4	21	91.3	19	Q90RF4	Q90rf4 human immun
5	21	91.3	19	Q90RH4	Q90rh4 human immun
6	21	91.3	26	Q6R4S9	Q6r4s9 pyrus commu
7	21	91.3	27	LSP_APTME	P09355 apis mellif
8	21	91.3	29	GALA_ONCMY	F47213 oncorhynch
9	21	91.3	29	SODC_OLEEU	P80740 olea europa
10	21	91.3	29	Q6X6Z8	Q6x6z8 enterococcu
11	21	91.3	29	Q6X701	Q6x701 enterococcu
12	21	91.3	29	Q8R398	Q8r398 mus musculu
13	21	91.3	30	Q8R9K8	Q8r9k8 rhizobium l
14	21	91.3	30	Q9JIC4	Q9jic4 rattus norv
15	21	91.3	31	Q9GKL4	Q9gkl4 canis famil
16	21	91.3	31	Q72DY1	Q72dy1 desulfovibr
17	21	91.3	32	Q7M553	Q7m553 halobacteri
18	21	91.3	32	Q6QLS6	Q6qls6 rhodococcus
19	21	91.3	32	Q736U2	Q736u2 bacillus ce
20	21	91.3	33	Q02832	Q02832 gorilla gor
21	21	91.3	33	Q66384	Q66384 thriobacillu
22	21	91.3	33	Q9F709	Q9f709 methylobaci
23	21	91.3	33	Q9F710	Q9f710 methylobacil
24	21	91.3	33	Q9F711	Q9f711 methylobact
25	21	91.3	34	HEMO_CHICK	P20057 gallus gall
26	21	91.3	34	Q73RE8	Q73re8 treponema d
27	21	91.3	35	Q6LBRO	Q6lbr0 mus musculu
28	21	91.3	35	Q8CLW4	Q8clw4 yersinia pe
29	21	91.3	36	Q8ZWQ4	Q8zwq4 pyrobaculum
30	21	91.3	36	Q9URE9	Q9ure9 saccharomyc
31	21	91.3	37	Q7TNQ0	Q7tnq0 mus musculu

32	21	91.3	38	2	Q6X700	Q6x700 enterococcu
33	21	91.3	38	2	Q8VJV6	Q8vjv6 mycobacteri
34	21	91.3	38	2	Q8ITC9	Q8itc9 bacillus an
35	21	91.3	38	2	Q9RUH4	Q9ruh4 deinococcus
36	21	91.3	39	2	Q6EES4	Q6ees4 homo sapien
37	21	91.3	39	2	Q9NC67	Q9nc67 dictyosteli
38	21	91.3	39	2	Q6X702	Q6x702 enterococcu
39	21	91.3	40	2	Q8ZTN8	Q8ztn8 pyrobaculum
40	21	91.3	40	2	P95768	P95768 staphylococ
41	21	91.3	40	2	Q48916	Q48916 mycobacteri
42	21	91.3	40	2	Q48917	Q48917 mycobacteri
43	21	91.3	40	2	Q48918	Q48918 mycobacteri
44	21	91.3	40	2	Q48951	Q48951 mycobacteri
45	21	91.3	40	2	Q49102	Q49102 mycobacteri

ALIGNMENTS

```
RESULT 1
Q80XX0 PRELIMINARY; PRT; 13 AA.
AC Q80XX0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypoxanthine phosphoribosyl transferase (Fragment).
GN Name=hprt;
OS Cricetus griseus (Chinese hamster).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;
OC Cricetus.
OX NCBI_TaxID=10029;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93024555; PubMed=1383700; DOI=10.1016/0027-5107(92)90198-B;
RA Fuscoe J.C., Zimmerman L.J., Fekete A., Setzer R.W., Rossiter B.J.;
RT 'Analysis of X-ray-induced HPRT mutations in CHO cells: insertion and
RT deletions.';
RL Mutat. Res. 269:171-183(1992).
DR EMBL; S46270; AAPI3884.1; -.
DR GO; GO:0016740; F:transferase activity; IEA.
KW Transferase.
FT NON_TER 13 13
SQ SEQUENCE 13 AA; 1516 MW; 113A46B08648F760 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 13;
Best Local Similarity 100.0%; Pred. No. 4e+02; 0; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

QY 3 PHG 5
DB 7 PHG 9

RESULT 2
O21922 PRELIMINARY; PRT; 16 AA.
AC O21922;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Integrase (Fragment).
GN Name=int;
OS Streptococcus thermophilus bacteriophage Sfi21.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Siphoviridae.
OX NCBI_TaxID=64186;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98008989; PubMed=9344917; DOI=10.1006/viro.1997.8769;
RA Bruttin A., Foley S., Brussow H.;
RT "The site-specific integration system of the temperate Streptococcus
RT thermophilus bacteriophage phisfi21.";
```

```
RL Virology 237:148-158(1997).
DR EMBL; AF013584; AAC48909.1; -.
DR EMBL; AF013587; AAC48910.1; -.
FT NON TER 1 1
SQ SEQUENCE 16 AA; 1856 MW; 8FA82D3270B9A959 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 4.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 14 PHG 16
RESULT 3
Q6LD80 PRELIMINARY; PRT; 18 AA.
AC Q6LD80;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Thromboipoietin receptor (Fragment).
GN Name=c-mpl1;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95166571; PubMed=7862460;
RA Alexander W.S., Dunn A.R.;
RT "Structure and transcription of the genomic locus encoding murine c-
RT Mpl, a receptor for thromboipoietin.";
RL Oncogene 10:795-803(1995).
DR EMBL; S76842; AAB33462.1; -.
DR GO; GO:0004872; F:receptor activity; IEA.
KW Receptor.
FT NON TER 18 18
SQ SEQUENCE 18 AA; 1808 MW; E01CB44EF83834E2 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 18;
Best Local Similarity 100.0%; Pred. No. 5.6e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 2 PHG 4
RESULT 4
Q90RF4 PRELIMINARY; PRT; 19 AA.
AC Q90RF4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=97CG282;
RX MEDLINE=21663216; PubMed=11804559; DOI=10.1089/08922202753394745;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., Obengui, M'Pandi M., M'Pele P., Harada Y.,
RA Ido E., Hayami M., Ichimura H., Parra H.Joseph.;
RT "Genetic subtypes of HIV type 1 based on the vpu/env sequences in the
RT Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 18:79-83(2002).
DR EMBL; AF127551; AAK84917.1; -.
FT NON TER 26 26
SQ SEQUENCE 26 AA; 3117 MW; 2A59C0F5BEA39D13 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 26;
DR EMBL; AF127551; AAK84917.1; -.
FT NON TER 1 1
SQ SEQUENCE 19 AA; 2163 MW; 2A83630189068236 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 7 PHG 9
RESULT 5
Q90RH4 PRELIMINARY; PRT; 19 AA.
AC Q90RH4;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Tat protein (Fragment).
GN Name=tat;
OS Human immunodeficiency virus 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=96CG17;
RX MEDLINE=21663216; PubMed=11804559; DOI=10.1089/08922202753394745;
RA Taniguchi Y., Takehsia J., Bikandou B., Mboudjeka I.,
RA N'Doundou-N'Kodia M.-Y., Obengui, M'Pandi M., M'Pele P., Harada Y.,
RA Ido E., Hayami M., Ichimura H., Parra H.Joseph.;
RT "Genetic subtypes of HIV type 1 based on the vpu/env sequences in the
RT Republic of Congo.";
RL AIDS Res. Hum. Retroviruses 18:79-83(2002).
DR EMBL; AF127546; AAK84900.1; -.
FT NON TER 1 1
SQ SEQUENCE 19 AA; 2180 MW; 2A81A56D59068236 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 19;
Best Local Similarity 100.0%; Pred. No. 5.9e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 7 PHG 9
RESULT 6
Q6R4S9 PRELIMINARY; PRT; 26 AA.
AC Q6R4S9;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE 1-aminocyclopropene-1-carboxylate synthase 1b (Fragment).
GN Name=DACS1b;
OS Pyrus communis (Pear).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids I; Rosales; Rosaceae; Maloideae; Pyrus.
OX NCBI_TaxID=23211;
RN [1]
RP SEQUENCE FROM N.A.
RA El Sharkawy I., Lelievre J.M., Latche A., Pech J.C.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DDeJ databases.
DR EMBL; AY514041; AAS66777.1; -.
DR HSSP; P37821; 1B8G.
FT NON TER 26 26
SQ SEQUENCE 26 AA; 3117 MW; 2A59C0F5BEA39D13 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 26;
```

Best Local Similarity 100.0%; Pred. No. 8.2e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 3; Conservative 0;

QY 3 PHG 5
Db 12 PHG 14

RESULT 7
LSP APIME STANDARD; PRT; 27 AA.
AC P09355;
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Larval-specific very high density lipoprotein (VHDL) (Fragment).
OS Apis mellifera (Honeybee).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;
OC Apidae;
OX NCBI_TaxID=7460;
RN [1]
RP SEQUENCE.
RC TISSUE=Hemolymph;
RX MEDLINE=87242376; PubMed=3109474;
RA Shipman B.A., Ryan R.O., Schmidt J.O., Law J.H.;
RT Purification and properties of a very high density lipoprotein from
the hemolymph of the honeybee *Apis mellifera*.;
RL Biochemistry 26:1885-1889(1987).
CC -!- FUNCTION: Unknown (it might play a role in lipid transport and/or
storage protein metabolism during metamorphosis).
CC -!- SUBUNIT: Homodimer.
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- TISSUE SPECIFICITY: Hemolymph.
CC -!- DEVELOPMENTAL STAGE: Present in high amounts in hemolymph only at
the end of larval life.
DR PIR; A27108; A27108.
KW Direct protein sequencing; Hemolymph; Lipoprotein.
FT NON_TER 27
SQ SEQUENCE 27 AA; 2885 MW; A9DDF7A75A65D10D CRC64;

Query Match 91.3%; Score 21; DB 1; Length 27;
Best Local Similarity 100.0%; Pred. No. 8.5e+02; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

QY 3 PHG 5
Db 5 PHG 7

RESULT 8
GALA ONCMY STANDARD; PRT; 29 AA.
ID AC P47213;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Galanin.
OS Oncorhynchus mykiss (Rainbow trout) (Salmo gairdneri).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei;
OC Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
OX NCBI_TaxID=8022;
RN [1]
RP SEQUENCE.
RC TISSUE=Stomach;
RX MEDLINE=95164756; PubMed=7532194;
RA Anglade I., Wang Y., Jensen J., Tramu G., Kah O., Conlon J.M.;
RT "Characterization of trout galanin and its distribution in trout brain
and pituitary".;
RL J. Comp. Neurol. 350:63-74(1994).
CC -!- FUNCTION: Contracts smooth muscle of the gastrointestinal and
genitourinary tracts, regulates growth hormone release, modulates

insulin release, and may be involved in the control of adrenal
secretion.
-!- SUBCELLULAR LOCATION: Secreted.
-!- SIMILARITY: Belongs to the galanin family.
DR InterPro; IPR008174; Galanin.
DR Pfam; PF01296; Galanin; 1.
DR PRINTS; PR00273; GALANIN.
DR ProDom; PD005962; Galanin; 1.
DR PROSITE; PS00861; GALANIN; 1.
KW Amidation; Direct protein sequencing; Hormone; Neuropeptide.
FT MOD_RES 29 Alanine amide.
SQ SEQUENCE 29 AA; 3044 MW; 73C37190403PA349 CRC64;

Query Match 91.3%; Score 21; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 9
SODC OLEEU STANDARD; PRT; 29 AA.
ID AC P80740;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Superoxide dismutase [Cu-Zn] (EC 1.15.1.1) (Allergen Ole e 5) (Ole e
V) (Fragment).
OS Olea europaea (Common olive).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Lamiales; Oleaceae; Olea.
OX NCBI_TaxID=4146;
RN [1]
RP SEQUENCE.
RC TISSUE=Pollen;
RX MEDLINE=98160390; PubMed=9500754;
RA Boluda L., Alonso C., Fernandez-Caldas E.;
RT "Purification, characterization, and partial sequencing of two new
allergens of Olea europaea".;
RL J. Allergy Clin. Immunol. 101:210-216(1998).
CC -!- FUNCTION: Destroys radicals which are normally produced within the
cells and which are toxic to biological systems (By similarity).
CC -!- CATALYTIC ACTIVITY: 2 superoxide + 2 H(+) = O(2) + H(2)O(2).
CC -!- COFACTOR: Binds 1 copper ion and 1 zinc ion per subunit (By
similarity).
CC -!- SIMILARITY: Belongs to the Cu-Zn superoxide dismutase family.
DR InterPro; IPR001424; SOD_CU_ZN.
DR Pfam; PF00080; Sod_Cu; 1.
DR PROSITE; PS00087; SOD_CU_ZN_1; PARTIAL.
DR PROSITE; PS00332; SOD_CU_ZN_2; PARTIAL.
KW Allergen; Antioxidant; Copper; Direct protein sequencing;
KW Metal-binding; Oxidoreductase; Zinc.
FT NON_TER 29
SQ SEQUENCE 29 AA; 2973 MW; 836C7A193EDAD71E CRC64;

Query Match 91.3%; Score 21; DB 1; Length 29;
Best Local Similarity 100.0%; Pred. No. 9.2e+02; Indels 0; Gaps 0;
Matches 3; Conservative 0; Mismatches 0;

QY 3 PHG 5
Db 13 PHG 15

RESULT 10
Q6X6Z8 PRELIMINARY; PRT; 29 AA.
ID Q6X6Z8
AC Q6X6Z8;
DT 05-JUL-2004 (TRENDEL. 27, Created)

Db 7 PHG 9
|||

RESULT 13

Q989K8 PRELIMINARY; PRT; 30 AA.
AC Q989K8;
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Mal6380 protein.
GN OrderedLocusNames=msl6380;
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFE303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
DR EMBL; AF003009; BAB52688.1; -.
KW Complete proteome.
SQ SEQUENCE 30 AA; 3291 MW; D966123BB111036A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 16 PHG 18

RESULT 14

Q9JIC4 PRELIMINARY; PRT; 30 AA.
AC Q9JIC4;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Sialomucin complex (Fragment).
GN Name=Muc4;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21060750; PubMed=10880365; DOI=10.1042/0264-6021:3490641;
RA Price-Schiavi S.A., Perez A., Barco R., Carraway K.L.;
RT "Cloning and characterization of the 5' flanking region of the
RT sialomucin complex/rat Muc4 gene: promoter activity in cultured
RT cells.";
RL Biochem. J. 349:641-649(2000).
DR EMBL; AF240632; AAF86958.1; -.
FT NON TER 30
SQ SEQUENCE 30 AA; 3355 MW; 94F779DCB226E2AD CRC64;

Query Match 91.3%; Score 21; DB 2; Length 30;
Best Local Similarity 100.0%; Pred. No. 9.5e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||

Db 4 PHG 6

RESULT 15

Q9GKL4 PRELIMINARY; PRT; 31 AA.
AC Q9GKL4;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Cardiac ryanodine receptor (Fragment).
GN Name=RYR2;
OS Canis familiaris (Dog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
OX NCBI_TaxID=9615;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21015404; PubMed=11130975;
RA Brouillette J.A., Andrew J.R., Venta P.J.;
RT "Estimate of nucleotide diversity in dogs with a pool-and-sequence
RT method.";
RL Mamm. Genome 11:1079-1086(2000).
DR EMBL; AF203021; AAG35635.1; -.
DR EMBL; AF203022; AAG35635.1; JOINED.
DR GO; GO:0004872; Fireceptor activity; IEA.
KW Receptor.
FT NON TER 1
FT NON TER 31
SQ SEQUENCE 31 AA; 3787 MW; 4E3B9EC795537E72 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 31;
Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
|||
Db 6 PHG 8

RESULT 16

Q72DY1 PRELIMINARY; PRT; 31 AA.
AC Q72DY1;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=DVU0798;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Daviden T.M., Zafar N., Zhou L., Radune D.,
RA Dimitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough.";
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AF017311; AAS95278.1; -.
DR TIGR; DVU0798; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 31 AA; 3438 MW; 0372F1F0B349FDE5 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 31;

```

Best Local Similarity 100.0%; Pred. No. 9.8e+02;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 18 PHG 20

RESULT 17
Q7M553 PRELIMINARY; PRT; 32 AA.
AC Q7M553;
DT 01-MAR-2004 (TRENBLrel. 26, Created)
DT 01-MAR-2004 (TRENBLrel. 26, Last sequence update)
DT 25-OCT-2004 (TRENBLrel. 28, Last annotation update)
DE Ribosomal protein S4 eR (Fragment)
OS Halobacterium salinarum (Halobacterium halobium).
OC Archaea; Euryarchaeota; Halobacteria; Halobacteriales;
OC Halobacteriaceae; Halobacterium.
OX NCBI_TaxID=2242;
RN [1]
RP SEQUENCE
RA Yaguchi M., Visentin L.P., Zuker M., Matheson A.T., Roy C.,
RA Strom A.R.;
RT "Amino-terminal sequences of ribosomal proteins from the 30S subunit
RT of archaeobacterium Halobacterium cutirubrum."
RL Submitted (DEC-1990) to the PIR data bank.
DR PIR; S11610; S11610. 1
FT NON_TER 1
FT NON_TER 32
SQ SEQUENCE 32 AA; 3507 MW; DC612B3DC957273A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 30 PHG 32

RESULT 18
Q6QLS6 PRELIMINARY; PRT; 32 AA.
AC Q6QLS6;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein (Fragment).
OS Rhodococcus fascians.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NRRL-B-15096;
RA Dorado G., Roldan J.M., Leshner J.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY534865; AAS45278.1; -.
KW Hypothetical protein.
FT NON_TER 1
FT NON_TER 32
SQ SEQUENCE 32 AA; 3463 MW; BCA73D0B5B1D338A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 26 PHG 28

us-10-074-225a-7.rup

RESULT 19
Q736U2 PRELIMINARY; PRT; 32 AA.
AC Q736U2;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS OrderedLocusNames=BCE2808;
GN Bacillus cereus (strain ATCC 10987).
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=222523;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14960714; DOI=10.1093/nar/gkh258;
RA Raeko D.A., Ravel J., Oekstad O.A., Helgason E., Cer R.Z., Jiang L.,
RA Shores K.A., Fouts D.E., Tourasse N.J., Angiuoli S.V., Kolonay J.F.,
RA Nelson W.C., Kolstoe A.-B., Fraser C.M., Read T.D.;
RT "The genome sequence of Bacillus cereus ATCC 10987 reveals metabolic
RT adaptations and a large plasmid related to Bacillus anthracis pXO1."
RL Nucleic Acids Res. 32:977-988(2004).
DR EMBL; AE017273; AAS41720.1; -.
DR TIGR; BCE2808; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 32 AA; 3635 MW; 6866784DFC0EAD3A3 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 32;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 14 PHG 16

RESULT 20
O02832 PRELIMINARY; PRT; 33 AA.
AC O02832;
DT 01-JUL-1997 (TRENBLrel. 04, Created)
DT 01-JUL-1997 (TRENBLrel. 04, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
DE Huntingtin protein (Fragment).
GN Name=IT15;
OS Gorilla Gorilla (Gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326790; PubMed=8766138;
RA Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
RT disease gene (it15) in several mammalian species."
RL Ann. Genet. 39:81-86(1996).
DR EMBL; S83377; AAB50771.1; -.
FT NON_TER 33
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 7 PHG 9

RESULT 21
O66384 PRELIMINARY; PRT; 33 AA.
ID O66384
AC O66384;
DT 01-AUG-1998 (TRENBLrel. 07, Created)

```

```

DT 01-AUG-1998 (Tremblrel. 07, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Beta subunit of light-harvesting I complex (Fragment).
GN Name=pufB;
OS Thiobacillus acidophilus.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhodospirillales;
OC Acetobacteraceae; Acidiphilium.
OX NCBI_TaxID=76588;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC27807T;
RX MEDLINE=99045881; PubMed=9828441;
RA Hiraishi A., Nagashima K.V.P., Matsuura K., Shimada K., Takaichi S.,
RA Wakao N., Katayama Y.;
RT "Phylogeny and photosynthetic features of Thiobacillus acidophilus and
RT related acidophilic bacteria: its transfer to the genus Acidiphilium
RT as Acidiphilium acidophilum comb. nov.";
RL Int. J. Syst. Bacteriol. 48:1389-1398(1998).
DR EMBL; AB013379; BAA26096.1; -.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3744 MW; 33A3B89C74B3E3BE CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 10 PHG 12

RESULT 22
Q9F709 PRELIMINARY; PRT; 33 AA.
AC Q9F709;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Gyrase subunit A (Fragment).
GN Name=gyrA;
OS Methylobacillus sp. SK1.
OC Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;
OC Methylophilaceae; Methylobacillus.
OX NCBI_TaxID=134932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SK1;
RA Kim H.J., Kim J.H., Kim K.S., Kim Y.M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF287970; AAG17771.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisom.
DR Pfam; PF00521; DNA_topoisom.
DR ProDom; PD000742; DNA_topoisom.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3791 MW; 32CC2C756A0FF45C CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

RESULT 23
Q9F710 PRELIMINARY; PRT; 33 AA.
AC Q9F710;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Gyrase subunit A (Fragment).
GN Name=gyrA;
OS Methylophilus methylotrophus (Bacterium W3A1).
OC Bacteria; Proteobacteria; Betaproteobacteria; Methylophilales;
OC Methylophilaceae; Methylophilus.
OX NCBI_TaxID=17;
RN [1]
RP SEQUENCE FROM N.A.
RA Kim H.J., Kim J.H., Kim K.S., Kim Y.M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF287969; AAG17770.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisom.
DR Pfam; PF00521; DNA_topoisom.
DR ProDom; PD000742; DNA_topoisom.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3791 MW; 32CC2C756A0FF45C CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

RESULT 24
Q9F711 PRELIMINARY; PRT; 33 AA.
AC Q9F711;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Gyrase subunit A (Fragment).
GN Name=gyrA;
OS Methylobacterium extorquens.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Methylobacteriaceae; Methylobacterium.
OX NCBI_TaxID=408;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AM1;
RA Kim H.J., Kim J.H., Kim K.S., Kim Y.M.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF287968; AAG17769.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisom.
DR Pfam; PF00521; DNA_topoisom.
DR ProDom; PD000742; DNA_topoisom.
FT NON_TER 1 1
FT NON_TER 33 33
SQ SEQUENCE 33 AA; 3791 MW; 32CC2C756A0FF45C CRC64;

Query Match 91.3%; Score 21; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 6 PHG 8

```

```
Db      ||||
        6 PHG 8

RESULT 25
HEMO_CHICK
ID HEMO_CHICK STANDARD; PRT; 34 AA.
AC P20057;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Hemopexin (Fragment).
GN Name=HPX;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE.
RX MEDLINE=68339942; PubMed=3421961;
RA Wallner D.; Cheng K.C.; Mueller-Eberhard U.;
RT "N-terminal amino acid sequences of the hemopexins from chicken, rat
and rabbit.";
RL Biochem. Biophys. Res. Commun. 155:622-625(1988).
CC -!- FUNCTION: Binds heme and transports it to the liver for breakdown
and iron recovery, after which the free hemopexin returns to the
circulation.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
DR PIR; C31514; C31514.
DR InterPro; IPR000585; Hemopexin.
DR PROSITE; PS00024; HEMOPEXIN; PARTIAL.
KW Direct protein sequencing; Glycoprotein; Heme; Plasma; Repeat;
KW Transport.
KW NON_TER
FT NON_TER 34 34
SQ SEQUENCE 34 AA; 3598 MW; 07A2EBAC27386C57 CRC64;

Query Match 91.3%; Score 21; DB 1; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db ||||
16 PHG 18

RESULT 26
Q73RE8
ID Q73RE8 PRELIMINARY; PRT; 34 AA.
AC Q73RE8;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=TDE0141;
OS Treponema denticola.
OC Bacteria; Spirochaetes; Spirochaetales; Spirochaetaceae; Treponema.
OX NCBI_TaxID=158;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=ATCC 35405 / DSM 14222;
RX PubMed=15064399; DOI=10.1073/pnas.0307639101;
RA Seshadri R.; Myers G.S.A.; Tettelin H.; Eisen J.A.; Heidelberg J.F.;
RA Dodson R.J.; Davidson T.M.; DeBoy R.T.; Fouts D.E.; Haft D.H.;
RA Selengut J.; Ren Q.; Brinkac L.M.; Madupu R.; Kolonay J.F.;
RA Durkin S.A.; Daugherty S.C.; Shetty J.; Shvartsbeyn A.;
RA Gregorogis E.; Geer K.; Teegaye G.; Malek J.A.; Ayodeji B.;
RA Shatsman S.; McLeod M.P.; Smajls D.; Howell J.K.; Pal S.; Amin A.;
RA Vashisth P.; McNeill T.Z.; Xiang Q.; Sodergren E.; Baca E.;
RA Weinstock G.M.; Norris S.J.; Fraser C.M.; Paulsen I.T.;
RT "Comparison of the genome of the oral pathogen Treponema denticola
with other spirochete genomes.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 101:5646-5651(2004).
DR EMBL; AE017246; AAS10639.1; -.
DR TIGR; TDE0141; -.
KW Complete proteome.
SQ SEQUENCE 34 AA; 3643 MW; EC84B7F1133531C4 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 34;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db ||||
28 PHG 30

RESULT 27
Q6LBR0
ID Q6LBR0 PRELIMINARY; PRT; 35 AA.
AC Q6LBR0;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE T10 class I MHC gene (exon 5) (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Domesticus;
RX MEDLINE=90063453; PubMed=2584927;
RA Broxson K.A.; Hunt S.W.; Hunkapiller T.; Sun H.Y.; Cheroutre H.;
RA Nickerson D.A.; Hood L.;
RT "Comparison of exon 5 sequences from 35 class I genes of the BALB/c
mouse.";
RL J. Exp. Med. 170:1837-1858(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=Domesticus;
RA Broxson K.;
RL Submitted (AUG-1989) to the EMBL/GenBank/DBJ databases.
DR EMBL; X16216; CAB82028.1; -.
FT NON_TER 1 1
FT NON_TER 35 35
SQ SEQUENCE 35 AA; 3770 MW; 695C1D11D6C7A052 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db ||||
8 PHG 10

RESULT 28
Q8CLW4
ID Q8CLW4 PRELIMINARY; PRT; 35 AA.
AC Q8CLW4;
DT 01-MAR-2003 (TReMBLrel. 23, Created)
DT 01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Hypothetical.
GN OrderedLocusNames=y0064;
OS Yersinia pestis.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=KIM5 / Biovar Mediaevalis;
RX MEDLINE=22137863; PubMed=12142430;
RX DOI=10.1128/JB.184.16.4601-4611.2002;
```

RA Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,
RA Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,
RA Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,
RA Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,
RA Perry R.D.;
RT "Genome sequence of *Yersinia pestis* KIM.";
RL J. Bacteriol. 184:4601-4611(2002).
DR EMBL: AE013607; AAM83659.1; -.
KW Hypothetical protein.
SQ SEQUENCE 35 AA; 3936 MW; 10744A0FDF6AC04B CRC64;
Query Match 91.3%; Score 21; DB 2; Length 35;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 33 PHG 35
RESULT 29
Q82WQ4 PRELIMINARY; PRT; 36 AA.
AC Q82WQ4;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FAREP2a.
GN OrderedLocusNames=PAE1672;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon *Pyrobaculum aerophilum*.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL: AB009833; AAL63646.1; -.
KW Complete proteome.
SQ SEQUENCE 36 AA; 4060 MW; 43CC1A624F2822C5 CRC64;
Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 10 PHG 12
RESULT 30
Q9URE9 PRELIMINARY; PRT; 36 AA.
AC Q9URE9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Multidrug resistance p-glycoprotein (Fragment).
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93368189; PubMed=8361216;
RA Allikmets R., Gerrard B., Stewart C., White M., Dean M.;
RA "Identification of P-glycoprotein/multidrug resistance genes from model organisms.";

RL Leukemia 7:13-13(1993).
CC -1- SIMILARITY: Belongs to the ABC transporter family.
DR GO: GO:0016020; Cmembrane; IEA.
DR GO: GO:0005524; F:ATP binding; IEA.
DR GO: GO:0042626; F:ATPase activity; IEA.
DR GO: GO:0006810; P:transport; IEA.
DR InterPro: IPR003439; ABC transporter.
DR Pfam: PF00005; ABC tran; 1.
DR ProDom: PD000006; ABC transporter; 1.
DR PROSITE: PS00211; ABC_TRANSPORTER_1; 1.
KW ATP-binding.
SQ SEQUENCE 36 AA; 3753 MW; CDB46C744DC8610D CRC64;
Query Match 91.3%; Score 21; DB 2; Length 36;
Best Local Similarity 100.0%; Pred. No. 1.1e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 2 PHG 4
RESULT 31
Q7TNQ0 PRELIMINARY; PRT; 37 AA.
AC Q7TNQ0;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RX MEDLINE=23288257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.B., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.P., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Udgin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6; TISSUE=Brain;
RA Strausberg R.;
RL Submitted (AUG-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC056177; AAH56177.1; -.
KW Hypothetical protein.
SQ SEQUENCE 37 AA; 3977 MW; 5D7166A450809A2B CRC64;
Query Match 91.3%; Score 21; DB 2; Length 37;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 PHG 5
Db 11

```
Db
RESULT 32
Q6X700
ID Q6X700 PRELIMINARY; PRT; 38 AA.
AC Q6X700;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Topoisomerase (Fragment).
GN Name=parC;
OS Enterococcus casseliflavus.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=37734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 25788;
RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;
RA Petersen A., Jensen L.B.;
RT "Analysis of gyrA and parC mutations in enterococci from environmental
RT samples with reduced susceptibility to ciprofloxacin.";
RL FEMS Microbiol. Lett. 231:73-76(2004).
DR EMBL; A252098; AAF84067.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0016853; F:isomerase activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
KW Isomerase.
FT NON_TER 1 1
FT NON_TER 38 38
SQ SEQUENCE 38 AA; 4357 MW; CEB751F1C22440E8 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 13 PHG 15

RESULT 33
Q8VJV6
ID Q8VJV6 PRELIMINARY; PRT; 38 AA.
AC Q8VJV6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=MT1909;
OS Mycobacterium tuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RX MEDLINE=2206494; PubMed=12218036;
DOI=10.1128/JB.184.19.5479-5490.2002;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J.D., DeBoy R.T., Dodson R.J., Gwinn M.L., Haft D.H.,
RA Hickey E.K., Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D.,
RA Salzberg S.L., Delcher A., Utterback T.R., Weidman J.F., Khouri H.M.,
RA Gill J., Mikula A., Bishai W., Jacobs W.R. Jr., Venter J.C.,
RA Fraser C.M.;
RT "Whole-genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains.";
```

```
RL J. Bacteriol. 184:5479-5490(2002).
DR EMBL; AE000516; AAK46180.1; -.
DR TIGR; MT1909; -.
KW Hypothetical protein.
SQ SEQUENCE 38 AA; 4273 MW; 461CADD7180DB359 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 10 PHG 12

RESULT 34
Q81TC9
ID Q81TC9 PRELIMINARY; PRT; 38 AA.
AC Q81TC9; Q6KVF5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BA1355, GBAA1355;
OS Bacillus anthracis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1392;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate Porton;
RX MEDLINE=22608414; PubMed=12721629; DOI=10.1038/nature01586;
RA Read T.D., Peterson S.N., Tourasse N.J., Baillie L.W., Paulsen I.T.,
RA Nelson K.E., Tettelin H., Fouts D.E., Eisen J.A., Gill S.R.,
RA Holtzapple E.K., Okstad O.A., Helgason E., Rilstone J., Wu M.,
RA Kolonay J.F., Beanan M.J., Dodson R.J., Brinkac L.M., Gwinn M.L.,
RA DeBoy R.T., Madpu R., Daugherty S.C., Durkin A.S., Haft D.H.,
RA Nelson W.C., Peterson J.D., Pop M., Khouri H.M., Radune D.,
RA Benton J.L., Mahamoud Y., Jiang L., Hance I.R., Weidman J.F.,
RA Berry K.J., Plaut R.D., Wolf A.M., Watkins K.L., Nierman W.C.,
RA Hazen A., Cline R.T., Redmond C., Thwaite J.E., White O.,
RA Salzberg S.L., Thomson B., Friedlander A.M., Koehler T.M.,
RA Hanna P.C., Kolstoe A.B., Fraser C.M.;
RT "The genome sequence of Bacillus anthracis Ames and comparison to
RT closely related bacteria.";
RL Nature 423:81-86(2003).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Ames / isolate 0581;
RA Ravel J., Rasko D.A., Shumway M.F., Jiang L., Cer R.Z., Federova N.B.,
RA Wilson M., Stanley S., Decker S., Read T.D., Salzberg S.L.,
RA Fraser C.M.;
RT "Bacillus anthracis comparative genomics.";
RL Submitted (MAY-2004) to the EMBL/GenBank/DBDJ databases.
DR EMBL; AE017028; AAP25301.1; -.
DR EMBL; AE017334; AAT30451.1; -.
DR TIGR; BA1355; -.
DR TIGR; GBAA1355; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 38 AA; 4456 MW; 965D6E445A9BE25C CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 21 PHG 23

RESULT 35
Q9RUH4
ID Q9RUH4 PRELIMINARY; PRT; 38 AA.
AC Q9RUH4;
```

```

DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Hypothetical protein DR1412.
GN OrderedLocNames=DR1412;
OS Deinococcus radiodurans.
OC Bacteria; Deinococcus-Thermus; Deinococci; Deinococcales;
OC Deinococcaceae; Deinococcus.
OX NCBI_TaxID=1299;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=R1 / ATCC 13939 / DSM 20539 / NCIB 9279;
RX MEDLINE=20036896; PubMed=10567266; DOI=10.1126/science.286.5444.1571;
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,
RA Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,
RA Vamathevan J.J., Lam P., McDonald L.A., Utterback T.R., Zalewski C.,
RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,
RA Ketchum K.A., Nelson K.E., Salzberg S.L., Smith H.O., Venter J.C.,
RA Fraser C.M.;
RT "Genome sequence of the radioresistant bacterium Deinococcus
RT radiodurans R1.";
RL Science 286:1571-1577(1999).
DR EMBL; AE001986; AAF10978.1; -.
DR PIR; G75398; G75398.
DR TIGR; DR1412; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 38 AA; 3891 MW; 0E8EFA6AFA325A4 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 38;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 31 PHG 33

RESULT 36
Q6EES4 PRELIMINARY; PRT; 39 AA.
AC Q6EES4;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE GRINL1A combined protein Gcom12 (Fragment).
GN Name=GRINL1A;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC Tissue=Lung;
RX PubMed=15233991; DOI=10.1016/j.ygeno.2004.04.004;
RA Roginski R.S., Mohan Raj B.K., Birditt B., Rowen L.;
RT "The human GRINL1A gene defines a complex transcription unit, an
RT unusual form of gene organization in eukaryotes.";
RL Genomics 84:265-276(2004).
DR EMBL; AY350748; AAQ82541.1; -.
FT NON_TER 1
SQ SEQUENCE 39 AA; 4393 MW; 530FE3E4CC04D4A5 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 22 PHG 24

RESULT 37

```

```

Q9NC67 PRELIMINARY; PRT; 39 AA.
AC Q9NC67;
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Sphingosine-1-phosphate lyase (Fragment).
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20431894; PubMed=10974109;
RA Li G., Alexander H., Schneider N., Alexander S.;
RT "Molecular basis for resistance to the anticancer drug cisplatin in
RT Dictyostelium.";
RL Microbiology 146:2219-2227(2000).
DR EMBL; AF233610; AAF97870.1; -.
DR DictyBase; DDB0214888; sglA.
DR GO; GO:0016829; F:lyase activity; IEA.
DR InterPro; IPR010916; TONB_Box_N.
DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; UNKNOWN_1.
KW Lyase.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4140 MW; BB528F43199727A CRC64;

Query Match 91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
DB 32 PHG 34

RESULT 38
Q6X702 PRELIMINARY; PRT; 39 AA.
AC Q6X702;
DT 05-JUL-2004 (Tremblrel. 27, Created)
DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)
DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)
DE Gyrase (Fragment).
GN Name=Gyrase;
OS Enterococcus faecalis.
OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.
OX NCBI_TaxID=37735;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CECT 4481;
RX PubMed=14769469; DOI=10.1016/S0378-1097(03)00929-7;
RA Petersen A., Jensen L.B.;
RT "Analysis of gyrase and parC mutations in enterococci from environmental
RT samples with reduced susceptibility to ciprofloxacin.";
RL FEMS Microbiol. Lett. 231:73-76(2004).
DR EMBL; AY252096; AAP84065.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisomIV.
DR Pfam; PF00521; DNA_topoisomIV; 1.
DR ProDom; PD000742; DNA_topoisomIV; 1.
FT NON_TER 1
FT NON_TER 39
SQ SEQUENCE 39 AA; 4312 MW; CDB3CDD0246265F3 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 39;
Best Local Similarity 100.0%; Pred. No. 1.2e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3 PHG 5
Db 31 PHG 33

RESULT 39

Q8ZTN8 PRELIMINARY; PRT; 40 AA.
AC Q8ZTN8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE FaREP2a.
GN OrderedLocusNames=PAE3166;
OS Pyrobaculum aerophilum.
OC Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;
OC Thermoproteaceae; Pyrobaculum.
OX NCBI_TaxID=13773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IM2 / ATCC 51768 / DSM 7523;
RX MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;
RA Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,
RA Miller J.H.;
RT "Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum
RT aerophilum.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).
DR EMBL; AB009915; AAL64721.1; -.
KW Complete proteome.
SQ SEQUENCE 40 AA; 4604 MW; B6C670022B6F64E7 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 14 PHG 16

RESULT 40

P95768 PRELIMINARY; PRT; 40 AA.
AC P95768;
DT 01-MAY-1997 (TrEMBLrel. 03, Created)
DT 01-MAY-1997 (TrEMBLrel. 03, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE GryA (Fragment).
GN Name=gryA;
OS Staphylococcus haemolyticus.
OC Bacteria; Firmicutes; Bacillales; Staphylococcus.
OX NCBI_TaxID=1283;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96254595; PubMed=8849234;
RA Yonezawa M., Takahata M., Banzawa-Futakuchi N., Matsubara N.,
RA Watanabe Y., Narita H., Matsunaga T., Igarashi H., Kawahara M.,
RA Onodera S.;
RT "DNA gyrase gryA mutations in quinolone-resistant clinical isolates of
RT Staphylococcus haemolyticus.";
RL Antimicrob. Agents Chemother. 40:1065-1066(1996).
DR EMBL; D78568; BAA11412.1; -.
DR HSSP; P09097; 1AB4.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0003918; F:DNA topoisomerase (ATP-hydrolyzing) activity; IEA.
DR GO; GO:0006265; P:DNA topological change; IEA.
DR InterPro; IPR002205; DNA_topoisomI.
DR Pfam; PF00521; DNA_topoisomI; 1.
DR ProDom; PD000742; DNA_topoisomI; 1.
FT NON_TER 1
FT NON_TER 40
SQ SEQUENCE 40 AA; 4533 MW; A4C6842027AA7812 CRC64;

Query Match 91.3%; Score 21; DB 2; Length 40;
Best Local Similarity 100.0%; Pred. No. 1.3e+03;
Matches 3; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 PHG 5
Db 13 PHG 15

Search completed: June 15, 2005, 14:21:22
Job time : 85.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225a-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseq1980s:*
2: Geneseq1990s:*
3: Geneseq2000s:*
4: Geneseq2001s:*
5: Geneseq2002s:*
6: Geneseq2003as:*
7: Geneseq2003bs:*
8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	5	ABB79809	Abb79809 Histidine
2	37	100.0	5	ADH10415	Adh10415 Rabbit HP
3	37	100.0	13	AAU10991	Aau10991 Cathepsin
4	37	100.0	33	AAE10907	Aae10907 Human bra
5	37	100.0	42	AAO07477	Aao07477 Human pol
6	37	100.0	97	AAG73663	Aag73663 Human col
7	37	100.0	101	ADH10412	Adh10412 Rabbit HP
8	37	100.0	137	ADM16706	Adm16706 PRM-PS-F1
9	37	100.0	148	ABB79807	Abb79807 Rabbit hi
10	37	100.0	148	ABB79806	Abb79806 Human his
11	37	100.0	148	ADH10411	Adh10411 Human HPR
12	37	100.0	151	AAW21476	Aaw21476 Peptide #
13	37	100.0	151	ABB43818	Abb43818 Peptide #
14	37	100.0	151	AAW37725	Aaw37725 Peptide #
15	37	100.0	151	ABB26753	Abb26753 Protein #
16	37	100.0	151	AAW77540	Aaw77540 Human bon
17	37	100.0	151	AAW64782	Aaw64782 Human bra
18	37	100.0	151	ABGS5189	Abg5189 Human liv
19	37	100.0	151	ABGA6584	Abg46584 Human pep
20	37	100.0	198	ABB71093	Abb71093 Drosophil
21	37	100.0	248	AAU55932	Aau55932 Propionib
22	37	100.0	248	ABW52451	Abw52451 Propionib
23	37	100.0	259	ABU35012	Abu35012 Protein e
24	37	100.0	264	ADL05482	Adl05482 M. catar
25	37	100.0	273	AAV32079	Aav32079 Human gap

26	37	100.0	273	5	ADI17057	Adi17057 Human NOV
27	37	100.0	273	5	ADI17060	Adi17060 Human NOV
28	37	100.0	273	6	ABR48232	Abr48232 Human bla
29	37	100.0	273	6	ABU56449	Abu56449 Lung canc
30	37	100.0	273	6	ABU56673	Abu56673 Lung canc
31	37	100.0	273	7	ADN39026	Adn39026 Cancer/an
32	37	100.0	273	7	ADN39932	Adn39932 Cancer/an
33	37	100.0	323	4	ABB69468	Abb69468 Drosophil
34	37	100.0	326	8	ADN46756	Adn46756 Thermococ
35	37	100.0	327	4	ABB66932	Abb66932 Drosophil
36	37	100.0	350	8	ADS30838	Ads30838 Bacterial
37	37	100.0	402	4	ABB71820	Abb71820 Drosophil
38	37	100.0	419	4	ABB59479	Abb59479 Drosophil
39	37	100.0	466	7	ABO77799	AbO77799 Pseudomon
40	37	100.0	525	5	ABB79804	Abb79804 Human his
41	37	100.0	525	8	ADE76897	Ade76897 Human pro
42	37	100.0	525	8	ADH10409	Adh10409 Human HPR
43	37	100.0	525	8	ADL12335	Adl12335 Human ste
44	37	100.0	526	5	ABB79805	Abb79805 Rabbit hi
45	37	100.0	526	8	ADH10410	Adh10410 Rabbit HP

ALIGNMENTS

RESULT 1

ABB79809
ID ABB79809 standard; peptide; 5 AA.

XX
AC ABB79809;

XX
DT 25-NOV-2002 (first entry)

XX
DE Histidine proline rich glycoprotein pentapeptide.

XX
KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW nontropic; neuroprotective; antiparkinsonian.

XX
OS Synthetic.

XX
PN WO200264621-A2.

XX
PD 22-AUG-2002.

XX
PF 14-FEB-2002; 2002WO-US004336.

XX
PR 14-FEB-2001; 2001US-0268370P.

XX
PA (ATTE-) ATTENUON LLC.

XX
PI Donate F, Harris S, Plunkett ML, Mazar AP;

XX
DR WPI; 2002-666989/71.

XX
DR P-PSDB; ABN84910.

XX
PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
diagnosing or treating diseases associated with undesired cell migration,
invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX
PS Claim 2; Page 67; 82pp; English.

XX
CC The present sequence is a specific example of claimed anti-angiogenic
pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
peptides comprise: the histidine-proline-rich (H/P) domain of human
histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
of rabbit HPRG (see ABB79807); a variant of these that is capable of
inhibiting angiogenesis, endothelial cell proliferation or endothelial
tube formation in vitro or in vivo; or a pentapeptide having the generic
sequence given in ABB79808, such as the present peptide, or its variant
having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
N- or C-terminus. Also claimed are: chemically synthesised or

CC recombinantly produced peptide multimers; a diagnostically or
 CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
 CC multimer; a diagnostically useful HPRG-related composition, comprising
 CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
 CC carrier; an antibody specific for an epitope of HPRG that is present in
 CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
 CC of the domains in a way which inhibits the anti-angiogenic activity of
 CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
 CC method for inhibiting cell migration, cell invasion, cell proliferation
 CC or angiogenesis, or for inducing apoptosis; a method for treating a
 CC subject having a disease or condition associated with undesired cell
 CC migration, invasion, proliferation, or angiogenesis; a method of
 CC stimulating or inhibiting angiogenesis in a subject; a method of
 CC detecting the presence of HPRG or its cleavage product or its peptide in
 CC a biological sample; isolated nucleic acids encoding the polypeptide,
 CC peptide or peptide multimer; an expression vector; transformed or
 CC transfected cells; a method of providing to a cell, tissue or organ an
 CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
 CC pentapeptide, or the peptide multimer; an affinity ligand useful for
 CC binding to, or isolating, an HPRG-binding molecule or cells expressing
 CC the binding molecule, comprising the polypeptide, peptide or peptide
 CC multimer, immobilised to a solid support or carrier; and a method of
 CC isolating HPRG-binding molecule, or isolating or enriching cells
 CC expressing HPRG-binding site or receptor, from a complex mixture. The
 CC compositions and methods are useful in diagnosing or treating a disease
 CC or condition associated with undesired cell migration, invasion,
 CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
 CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
 CC fractures. HPRG is especially useful in inhibiting the growth of primary
 CC tumours or metastases, and may also be used in treating neurodegenerative
 CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays

XX Sequence 5 AA;

Query Match 100.0%; Score 37; DB 5; Length 5;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 1 HHPHG 5

RESULT 2

ID ADH10415

XX ADH10415 standard; peptide; 5 AA.

AC ADH10415;

XX 11-MAR-2004 (first entry)

DE Rabbit HPRG protein H/P rich domain repeat fragment.

XX Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipsoriatic; dermatological; cardiac; vasotropic; vulnery;

XX angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

OS WO2003077872-A2.

PN 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

PF 15-MAR-2002; 2002US-0364047P.

XX (ATTB-) ATTENUON LLC.

XX Mccrae K, Donate F, Juarez J, Mazar AP;
 PI WPI; 2004-090604/09.
 DR New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 XX inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.

XX Example 5; SEQ ID NO 29; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
 CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a repeat
 CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 100.0%; Score 37; DB 8; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 1 HHPHG 5

RESULT 3

AAU10991

XX AAU10991 standard; peptide; 13 AA.

AC AAU10991;

XX 26-MAR-2002 (first entry)

XX Cathepsin D substrate.

XX Contrast agent; cancer; cardiovascular disease; inflammation; stroke;
 KW central nervous system disorder; infection; cardiac failure; embolism;
 KW myocardial infarction; atherosclerosis; thrombosis; aneurysm;
 KW haemorrhage; Alzheimer's disease; multiple sclerosis; bone disease;
 KW osteoporosis; apoptosis; necrosis; cathepsin D.

OS Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1

FT /note= "Pn216-succinyl. Pn216 = Bis[(1,1-dimethyl-2-N-
 FT hydroxyimino propyl) aminoethyl]-2-aminoethyl amine"

FT Modified-site 9

FT	/note= "Cys(Et) "	
XX		
PN	WO200189584-A2.	
XX		
PD	29-NOV-2001.	
XX		
PF	23-MAY-2001; 2001WO-NO000215.	
XX		
PR	23-MAY-2000; 2000NO-00002644.	
XX		
PA	(NYCO-) NYCOMED IMAGING AS.	
XX		
PI	Klaveness J, Tolleshaug H;	
XX		
DR	WPI; 2002-097616/13.	
XX		
PT	Contrast agents susceptible of changing pharmacodynamic and/or	
PT	pharmacokinetic properties upon enzymatic activity, useful in diagnosis	
PT	of cancer, cardiovascular diseases and inflammation in humans and	
PT	animals.	
XX		
PS	Example 2; Page 53; 77pp; English.	
XX		
CC	The invention relates to a contrast agent substrate (I) susceptible of	
CC	changing pharmacodynamic and/or pharmacokinetic properties upon the	
CC	influence of enzymatic activity. The contrast agent substrate is useful	
CC	for detection of tissue or cells with abnormal metabolic activity, for	
CC	identification and/or diagnosis of cancer, cardiovascular diseases,	
CC	diseases on the central nervous system, inflammations, or infections and	
CC	detection of an area of a disease of abnormal enzymatic activity, where	
CC	the substrate is administered to human or animal body and a contrast	
CC	agent signal is detected as a result of the contrast agent changing	
CC	pharmacodynamic and/or pharmacokinetic properties. The substrate is	
CC	useful for manufacturing a medicament for detecting an area of disease of	
CC	abnormal enzymatic activity. The contrast agent is useful in diagnosis of	
CC	cardiac failure, myocardial infarction, atherosclerosis, thrombosis,	
CC	embolism, aneurysm, stroke, haemorrhage, central nervous system	
CC	diseases, preferably Alzheimer's disease or multiple sclerosis, bone	
CC	diseases such as osteoporosis, viral infections, and for identification	
CC	of apoptosis and necrosis. The metabolically sensitive contrast agents	
CC	are more sensitive to pathology than morphological contrast agents. As	
CC	abnormal enzymatic activity is an early sign of disease/condition, the	
CC	contrast agents have a potential for diagnosing disease at an early	
CC	stage, which in many clinical situations are important for the outcome of	
CC	the treatment. The agents are very sensitive to treatment and can be used	
CC	to follow up treatment. The present sequence represents the cathepsin D	
XX	substrate, used in the method of the invention	
SQ	Sequence 13 AA;	
	Query Match 100.0%; Score 37; DB 5; Length 13;	
	Best Local Similarity 100.0%; Pred. No. 4.8;	
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HHPHG 5	
DB		
	2 HHPHG 6	
RESULT 4		
AAE10907		
ID	AAE10907 standard; peptide; 33 AA.	
XX		
AC	AAE10907;	
XX		
DT	18-DEC-2001 (first entry)	
XX		
DE	Human brain sodium channel 4 protein (BNAc4) C-terminal peptide.	
XX		
KW	Human; Brain sodium channel 4 protein; BNAc4; cerebroprotective;	
KW	acid-sensing ion channel; ASIC4; neuromodulation; vasotropic;	
KW	neurotransmission; ischaemia; neurodegeneration; Alzheimer's disease;	
KW	Huntington's disease; amyotrophic lateral sclerosis; antiinflammatory;	
KW	pain; Parkinsonism; inflammatory condition; cerebellar ataxia;	
KW	complex regional pain syndrome; osteoarthritis; rheumatoid arthritis;	
KW	ulcer; neurodegenerative disease; asthma; irritable bowel syndrome;	
KW	chronic obstructive pulmonary disease; central nervous system disease;	
KW	psoriasis; intestinal tract disease; gene therapy; neuroprotective;	
XX	neotropic; analgesic; osteopathic; cancer.	
OS	Homo sapiens.	
XX		
FH	Key Location/Qualifiers	
FT	Misc-difference 2 /label= Lys, Arg	
FT	Misc-difference 5 /label= Ala, Val	
FT		
XX	WO200166125-A1.	
XX		
XX	13-SEP-2001.	
PD		
PF	13-FEB-2001; 2001WO-US004525.	
XX		
PR	03-MAR-2000; 2000US-00518959.	
XX		
PA	(ORTH) ORTHO-MCNEIL PHARM INC.	
XX		
PI	Dubin AE, Erlander MG, Huvar R, Pyati J;	
XX		
DR	WPI; 2001-589913/66.	
XX		
PT	Novel human brain sodium channel 4 protein, a member of acid-sensing ion	
PT	channel family, for identifying modulators comprises measuring the effect	
PT	of the modulator on the protein.	
XX		
PS	Example 7; Page 60; 94pp; English.	
XX		
CC	The invention relates to human brain sodium channel 4 protein (BNAc4), a	
CC	member of acid-sensing ion channel family (ASIC4). Human BNAc4 is useful	
CC	for identifying compounds that modulate its activity, by measuring the	
CC	effect of the modulator on the protein, BNAc4-dependent cytotoxicity or	
CC	BNAc4-dependent activation of ion flux. Compounds discovered by screening	
CC	test compounds using human BNAc4 protein is useful for treating a patient	
CC	in need of a treatment for a condition mediated by human BNAc4. BNAc4	
CC	antibody is useful for blocking the activity of the polypeptide. BNAc4	
CC	plays an important role in neuromodulation, neurotransmission, pain,	
CC	ischaemia and neurodegeneration underlying diseases including Alzheimer's	
CC	disease, Huntington's disease, amyotrophic lateral sclerosis, cerebellar	
CC	ataxias and Parkinsonism. Modulators identified using this protein are	
CC	useful as therapeutic agents, for the treatment of ischaemia inflammatory	
CC	conditions and for use as analgesics for intractable pain, complex	
CC	regional pain syndrome, arthritis (osteoarthritis, rheumatoid), ulcers,	
CC	neurodegenerative diseases, asthma, chronic obstructive pulmonary	
CC	disease, irritable bowel syndrome, psoriasis, central nervous system	
CC	diseases, diseases of the intestinal tract, abnormal proliferation and	
CC	cancer especially in the digestive system and female gonads, ulcer, liver	
CC	disease, control of viscera innervated by the dorsal root ganglia and to	
CC	diagnose or treat any disorder related to abnormal expression of the	
CC	human BNAc4 polypeptides. BNAc4 DNA and its portions are useful for	
CC	isolating homologues of the DNA molecules, identifying and isolating	
CC	genomic equivalents of the DNA molecules and identifying, detecting or	
CC	isolating mutant forms of the DNA molecules. BNAc4 DNA is used in gene	
CC	therapy. The present sequence is human BNAc4 C-terminal peptide. There	
CC	were 2 bases in the BNAc4 coding sequence that were consistently	
CC	different in one of 10 independent cDNA library clones. These keto and	
CC	pyrimidine differences caused an L to become R and A to V in the C-	
CC	terminal peptide	
XX		
SQ	Sequence 33 AA;	
	Query Match 100.0%; Score 37; DB 4; Length 33;	
	Best Local Similarity 100.0%; Pred. No. 13;	
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 HHPHG 5	

```

Db          |||||
            18 HHPHG 22

RESULT 5
ID AAO07477 standard; protein; 42 AA.
XX
XX AAO07477;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX
XX Human polypeptide SEQ ID NO 21369.
DE
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
XX Homo sapiens.
OS
XX
XX WO200164835-A2.
PN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PF
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX
XX 18-MAY-2000; 2000US-00577409.
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI; 2001-514838/56.
DR
XX
XX N-PSDB; AAI87408.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
XX Claim 20; SEQ ID NO 21369; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 42 AA;
SQ
Query Match 100.0%; Score 37; DB 4; Length 42;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
Db 38 HHPHG 42

RESULT 6
ID AAG73663 standard; protein; 97 AA.
XX
XX AAG73663;
AC
XX
XX AAG73663;
KW
XX

```

```

DT 03-SEP-2001 (first entry)
XX
XX Human colon cancer antigen protein SEQ ID NO:4427.
XX
XX Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW colorectal carcinoma; chromosome 1.
XX
XX Homo sapiens.
OS
XX
XX WO200122920-A2.
PN
XX
XX 05-APR-2001.
PD
XX
XX 28-SEP-2000; 2000WO-US026524.
PF
XX
XX 29-SEP-1999; 99US-0157137P.
PR
XX
XX 03-NOV-1999; 99US-0163280P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
PA
XX
XX Ruben SM, Barash SC, Birse CE, Rosen CA;
PI
XX
XX WPI; 2001-235357/24.
DR
XX
XX N-PSDB; AAH33094.
XX
XX Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT useful for preventing, diagnosing and/or treating colorectal cancers.
XX
XX Claim 11; Page 6244-6245; 9803pp; English.
XX
XX AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon
CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
CC proteins are collectively known as colon cancer antigens. The colon
CC cancer antigens have cytostatic activity and can be used in gene therapy
CC and vaccine production. N and P may be used in the prevention, diagnosis
CC and treatment of diseases associated with inappropriate P expression. For
CC example, N and P may be used to treat disorders associated with decreased
CC expression by rectifying mutations or deletions in a patient's genome
CC that affect the activity of P by expressing inactive proteins or to
CC supplement the patients own production of P. Additionally, N may be used
CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the proteins. N and P
CC can be used in the prevention, diagnosis and treatment of colorectal
CC carcinomas and cancers. AAH37196 to AAH37204 and AAG77789 represent
CC sequences used in the exemplification of the present invention. N.B.
CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
CC time of publication, meaning no sequences are present for SEQ ID NO:1027
CC to 1052, 7921 and 7922
XX
XX Sequence 97 AA;
SQ
Query Match 100.0%; Score 37; DB 4; Length 97;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
Db 53 HHPHG 57

RESULT 7
ID ADH10412 standard; protein; 101 AA.
XX
XX ADH10412;
AC
XX
XX 11-MAR-2004 (first entry)
DT
XX
XX Rabbit HPRG protein H/P rich domain.
XX
XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;

```

KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
 KW angiogenesis; gene therapy; rabbit.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO2003077872-A2.
 XX
 XX 25-SEP-2003.
 PD
 XX
 PF 17-MAR-2003; 2003WO-US008060.
 XX
 XX 15-MAR-2002; 2002US-0364047P.
 PR
 XX (ATTE-) ATTENUON LLC.
 PA
 XX

PI Mcrae K, Donate F, Juarez J, Mazar AP;
 XX WPI; 2004-090604/09.
 DR
 XX

XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 XX

PS Claim 6; SEQ ID NO 24; 117pp; English.
 XX

CC The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
 CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a rabbit HPRG
 CC protein His-Pro (H/P) rich domain.
 XX

SQ Sequence 101 AA;
 XX

Query Match 100.0%; Score 37; DB 8; Length 101;
 Best Local Similarity 100.0%; Pred. No. 39;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
 |||||
 Db 13 HHPHG 17

RESULT 8
 ADM16706

ID ADM16706 standard; protein; 137 AA.
 XX

XX ADM16706;
 AC

XX 17-JUN-2004 (first entry)
 DT

XX PRM-P5-F12 protein.
 DE

XX
 KW

Phlebotomus ariasi; Phlebotomus perniciosus; Vaccine;
 Leishmania infection.

XX
 OS

Phlebotomus ariasi.

XX
 PN

WO2004027041-A2.

XX
 PD

01-APR-2004.

XX
 PF

18-SEP-2003; 2003WO-US029833.

XX
 PR

19-SEP-2002; 2002US-0412327P.

XX
 PR

12-NOV-2002; 2002US-0425852P.

XX
 PA

(MERI-) MERIAL LTD.

XX
 PA

(USSH) US DEPT HEALTH & HUMAN SERVICES.

XX
 PI

Valenzuela JG, Ribiero JMC, Kamhawi S, Belkaid Y, Fischer L;

XX
 PI

Audonnet J, Milward F;

XX
 DR

WPI; 2004-295410/27.

XX
 DR

N-PSDB; ADM16707.

XX
 PT

New substantially purified salivary Phlebotomus ariasi or Phlebotomus

XX
 PT

perniciosus polypeptide, useful for inhibiting, treating or preventing

XX
 PS

Claim 1; SEQ ID NO 29; 200pp; English.

XX
 CC

The present invention relates to a substantially purified salivary

XX
 CC

Phlebotomus ariasi or Phlebotomus perniciosus polypeptide. The

XX
 CC

composition comprising the polypeptide or the nucleic acid encoding the

XX
 CC

polypeptide is useful for manufacture of a medicament. The polypeptides

XX
 CC

and nucleic acids are useful for inducing an immune response to a P.

XX
 CC

ariasi or P. perniciosus polypeptide. It is also useful for inhibiting a

XX
 CC

symptom of a Leishmania infection or preventing a Leishmania infection in

XX
 CC

a subject. The present sequence represents a purified Phlebotomus ariasi

XX
 SQ

Sequence 137 AA;

Query Match 100.0%; Score 37; DB 8; Length 137;
 Best Local Similarity 100.0%; Pred. No. 54;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
 |||||
 Db 61 HHPHG 65

RESULT 9
 ABB79807

ID ABB79807 standard; protein; 148 AA.

XX
 AC

ABB79807;

XX
 XX

25-NOV-2002 (first entry)

XX
 DE

Rabbit histidine proline rich glycoprotein His/Pro-rich domain.

XX
 XX

Histidine proline rich glycoprotein; HPGP; rabbit; antiangiogenic;

XX
 KW

cytostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;

XX
 KW

gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;

XX
 KW

ophthalmological; neurotropic; neuroprotective; antiparkinsonian; chelator.

OS
 Oryctolagus cuniculus.

XX
 XX

WO200264621-A2.

XX
 PN

22-AUG-2002.

XX
 PD

14-FEB-2002; 2002WO-US004336.

XX
 PF

XX PR 14-FEB-2001; 2001US-0268370P.
XX PA (ATTE-) ATTENUON LLC.
XX PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX DR WPI; 2002-666989/71.
XX DR P-PSDB; ABN84910.
XX DR New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX PS
XX Claim 1; Page 13; 82pp; English.
XX CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of rabbit anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79805). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain (see ABB79806) of human
CC HPRG; an H/P domain of rabbit HPRG; a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
XX immunoassays
XX SQ Sequence 148 AA;
Query Match 100.0%; Score 37; DB 5; Length 148;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHHPG 5
|||
Db 35 HHHPG 39
RESULT 10
ABB79806
ID ABB79806 standard; protein; 148 AA.

XX AC ABB79806;
XX DT 25-NOV-2002 (first entry)
XX DE Human histidine proline rich glycoprotein His/Pro-rich domain.
XX KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cystostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian.
XX OS Homo sapiens.
XX PN WO200264621-A2.
XX PD 22-AUG-2002.
XX PF 14-FEB-2002; 2002WO-US004336.
XX PR 14-FEB-2001; 2001US-0268370P.
XX PA (ATTE-) ATTENUON LLC.
XX PI Donate F, Harris S, Plunkett ML, Mazar AP;
XX DR WPI; 2002-666989/71.
XX DR P-PSDB; ABN84910.
XX PT New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX PS Claim 1; Page 13; 82pp; English.
XX CC The present sequence is the protein sequence of the histidine-proline-
CC rich (H/P) domain of human anti-angiogenic histidine proline rich
CC glycoprotein (HPRG, see also ABB79804). Claimed anti-angiogenic
CC polypeptides or peptides comprise: the H/P domain of human HPRG; an H/P
CC domain of rabbit HPRG (see ABB12345); a variant of these that is capable
CC of inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or
CC therapeutically labeled anti-angiogenic polypeptide, peptide or peptide
CC multimer; a diagnostically useful HPRG-related composition, comprising
CC the diagnostically labeled polypeptide, peptide or peptide multimer and a
CC carrier; an antibody specific for an epitope of HPRG that is present in
CC the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any
CC of the domains in a way which inhibits the anti-angiogenic activity of
CC HPRG or the domain, or an antigen-binding fragment of the antibody; a
CC method for inhibiting cell migration, cell invasion, cell proliferation
CC or angiogenesis, or for inducing apoptosis; a method for treating a
CC subject having a disease or condition associated with undesired cell
CC migration, invasion, proliferation, or angiogenesis; a method of
CC stimulating or inhibiting angiogenesis in a subject; a method of
CC detecting the presence of HPRG or its cleavage product or its peptide in
CC a biological sample; isolated nucleic acids encoding the polypeptide,
CC peptide or peptide multimer; an expression vector; transformed or
CC transfected cells; a method of providing to a cell, tissue or organ an
CC angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its
CC pentapeptide, or the peptide multimer; an affinity ligand useful for
CC binding to, or isolating, an HPRG-binding molecule or cells expressing
CC the binding molecule, comprising the polypeptide, peptide or peptide
CC multimer, immobilised to a solid support or carrier; and a method of
CC isolating HPRG-binding molecule, or isolating or enriching cells
CC expressing HPRG-binding site or receptor, from a complex mixture. The
CC compositions and methods are useful in diagnosing or treating a disease
CC or condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic
CC retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or
CC fractures. HPRG is especially useful in inhibiting the growth of primary
CC tumours or metastases, and may also be used in treating neurodegenerative
CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
CC stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
XX immunoassays

CC diseases like Alzheimer's or Parkinson's disease. The antibodies are
 CC stimulators of angiogenesis and are useful for promoting
 CC neovascularization in pertinent disease states, and in various
 CC immunoassays

XX SQ Sequence 148 AA;
 Query Match 100.0%; Score 37; DB 5; Length 148;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 Db 35 HHPHG 39

RESULT 11
 ADH10411
 ID ADH10411 standard; protein; 148 AA.
 XX AC ADH10411;
 XX DT 11-MAR-2004 (first entry)
 XX DE Human HPRG protein H/P rich domain.
 XX KW Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipsoriatic; dermatological; cardiac; vasotropic; vulnary;
 KW angiogenesis; gene therapy; human.
 XX OS Homo sapiens.
 XX PN WO2003077872-A2.
 XX PD 25-SEP-2003.
 XX PF 17-MAR-2003; 2003WO-US008060.
 XX PR 15-MAR-2002; 2002US-0364047P.
 XX PA (ATTE-) ATTENUON LLC.
 XX PI Mcrae K, Donate F, Juarez J, Mazar AP;
 XX WPI; 2004-090604/09.
 XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 XX Claim 6; SEQ ID NO 23; 117pp; English.

CC The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
 CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing

CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a human HPRG.
 CC protein His-Pro (H/P) rich domain.
 XX SQ Sequence 148 AA;
 Query Match 100.0%; Score 37; DB 8; Length 148;
 Best Local Similarity 100.0%; Pred. No. 58;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 Db 35 HHPHG 39

RESULT 12
 AAM21476
 ID AAM21476 standard; protein; 151 AA.
 XX AC AAM21476;
 XX DT 12-OCT-2001 (first entry)
 XX DE Peptide #7910 encoded by probe for measuring cervical gene expression.
 XX KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX OS Homo sapiens.
 XX PN WO200157278-A2.
 XX PD 09-AUG-2001.
 XX PF 30-JAN-2001; 2001WO-US000670.
 XX PR 04-FEB-2000; 2000US-0180312P.
 XX PR 26-MAY-2000; 2000US-0207456P.
 XX PR 30-JUN-2000; 2000US-00608408.
 XX PR 03-AUG-2000; 2000US-00632366.
 XX PR 21-SEP-2000; 2000US-0234687P.
 XX PR 27-SEP-2000; 2000US-0236359P.
 XX PR 04-OCT-2000; 2000GB-00024263.
 XX PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488901/53.
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human cervical epithelial cells.
 XX Claim 27; SEQ ID NO 26302; 487pp; English.

CC The present invention relates to human single exon nucleic acid probes
 CC (SENP; see A110068-A128459). The present sequence is a peptide encoded
 CC by one such probe. The SENPs are derived from human HeLa cells. The SENPs
 CC can be used to produce a single exon microarray, which can be used for
 CC measuring human gene expression in a sample derived from human cervical
 CC epithelial cells. By measuring gene expression, the probes are therefore
 CC useful in grading and/or staging of diseases of the cervix, notably
 CC cervical cancer. Note: The sequence data for this patent did not form
 CC part of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ffp.wipo.int/pub/published_pct_sequences
 XX Sequence 151 AA;
 SQ

```
Query Match      100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      128 HHPHG 132

RESULT 13
ABB43818
ID ABB43818 standard; peptide; 151 AA.
XX
XX AC ABB43818;
XX
XX DT 04-FEB-2002 (first entry)
XX
DE Peptide #11324 encoded by human foetal liver single exon probe.
XX
XX KW Human; foetal liver; gene expression; single exon nucleic acid probe.
XX
XX OS Homo sapiens.
XX
XX PN WO200157277-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000669.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-00632366.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX PF WPI; 2001-488897/52.
XX
XX PR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human foetal liver.
XX
XX PS Claim 27; SEQ ID NO 36453; 639pp + Sequence Listing; English.
XX
XX CC The invention relates to a single exon nucleic acid probe for measuring
XX CC human gene expression in a sample derived from human foetal liver. The
XX CC single exon nucleic acid probes may be used for predicting, measuring and
XX CC displaying gene expression in samples derived from human foetal liver. The
XX CC present sequence is a peptide encoded by a single exon nucleic acid probe
XX CC of the invention. Note: The sequence data for this patent did not form
XX CC part of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 151 AA;

Query Match      100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      128 HHPHG 132

RESULT 14
AAM37725
ID AAM37725 standard; protein; 151 AA.
XX
XX AC AAM37725;

Query Match      100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      128 HHPHG 132

RESULT 15
ABB26753
ID ABB26753 standard; protein; 151 AA.
XX
XX AC ABB26753;
XX
XX DT 23-JAN-2002 (first entry)
XX
XX DE Protein #8752 encoded by probe for measuring heart cell gene expression.
XX
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200157274-A2.
XX
XX PD 09-AUG-2001.
XX
```

```
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #11762 encoded by probe for measuring placental gene expression.
XX
XX KW Probe; microarray; human; placenta; antenatal diagnosis;
XX KW genetic disorder.
XX
XX OS Homo sapiens.
XX
XX PN WO200157272-A2.
XX
XX PD 09-AUG-2001.
XX
XX PF 30-JAN-2001; 2001WO-US000663.
XX
XX PR 04-FEB-2000; 2000US-0180312P.
XX
XX PR 26-MAY-2000; 2000US-0207456P.
XX
XX PR 30-JUN-2000; 2000US-00608408.
XX
XX PR 03-AUG-2000; 2000US-00632366.
XX
XX PR 21-SEP-2000; 2000US-0234687P.
XX
XX PR 27-SEP-2000; 2000US-0236359P.
XX
XX PR 04-OCT-2000; 2000GB-00024263.
XX
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
XX PF WPI; 2001-488897/53.
XX
XX PR Human genome-derived single exon nucleic acid probes useful for analyzing
XX PT gene expression in human placenta.
XX
XX PS Claim 27; SEQ ID NO 37994; 654pp; English.
XX
XX CC The present invention relates to single exon nucleic acid probes (SENP:
XX CC see AA1315-AA157546). The present sequence is a peptide encoded by one
XX CC such probe. The probes are useful for producing a microarray for
XX CC predicting, measuring and displaying gene expression in samples derived
XX CC from human placenta. The probes are useful for antenatal diagnosis of
XX CC human genetic disorders
XX
XX SQ Sequence 151 AA;

Query Match      100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      128 HHPHG 132

RESULT 15
ABB26753
ID ABB26753 standard; protein; 151 AA.
XX
XX AC ABB26753;
XX
XX DT 23-JAN-2002 (first entry)
XX
XX DE Protein #8752 encoded by probe for measuring heart cell gene expression.
XX
XX KW Human; gene expression; heart; microarray; vascular system;
XX KW cardiovascular disease; hypertension; cardiac arrhythmia;
XX KW congenital heart disease.
XX
XX OS Homo sapiens.
XX
XX PN WO200157274-A2.
XX
XX PD 09-AUG-2001.
XX
```

```

PF 30-JAN-2001; 2001WO-US000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488899/53.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
PS Claim 15; SEQ ID NO 28523; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 128 HHPHG 132

RESULT 16
AAM77540
ID AAM77540 standard; protein; 151 AA.
XX
AC AAM77540;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human bone marrow expressed probe encoded protein SEQ ID NO: 37846.
XX
DE Human bone marrow expressed exon; gene expression analysis; probe;
XX KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX KW
XX OS Homo sapiens.
XX OS
XX PN WO200157276-A2.
XX PN
XX PD 09-AUG-2001.
XX PD
XX PF 30-JAN-2001; 2001WO-US000668.
XX PF
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR
XX PR 30-JUN-2000; 2000US-00608408.
XX PR
XX PR 03-AUG-2000; 2000US-00632366.
XX PR
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR
XX PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 36887; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,

```

```

XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-488900/53.
XX
PT Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human bone marrow.
XX
PS Example 4; SEQ ID NO 37846; 658pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC bone marrow. They can be used to measure gene expression in bone marrow
XX CC samples, which may enable the improved diagnosis and treatment of cancers
XX CC such as lymphoma, leukaemia and myeloma. The present sequence is a
XX CC protein encoded by one of the probes of the invention
XX
SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 128 HHPHG 132

RESULT 17
AAM64782
ID AAM64782 standard; protein; 151 AA.
XX
AC AAM64782;
XX
DT 05-NOV-2001 (first entry)
XX
DE Human brain expressed single exon probe encoded protein SEQ ID NO: 36887.
XX
DE Human; brain expressed exon; gene expression analysis; probe; microarray;
XX KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX KW
XX OS Homo sapiens.
XX OS
XX PN WO200157275-A2.
XX PN
XX PD 09-AUG-2001.
XX PD
XX PF 30-JAN-2001; 2001WO-US000667.
XX PF
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR
XX PR 30-JUN-2000; 2000US-00608408.
XX PR
XX PR 03-AUG-2000; 2000US-00632366.
XX PR
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR
XX PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX
DR WPI; 2001-483446/52.
XX
PT Single exon nucleic acid probes for analyzing gene expression in human
PT brains.
XX
PS Example 4; SEQ ID NO 36887; 650pp + Sequence Listing; English.
XX
CC The present invention provides a number of single exon nucleic acid
XX CC probes which are derived from genomic sequences expressed in the human
XX CC brain. They can be used to measure gene expression in brain cell samples,

```

CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention

XX SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 128 HHPHG 132

RESULT 18
 ABG59189
 ID ABG59189 standard; peptide; 151 AA.

XX AC ABG59189;
 XX DT 25-FEB-2003 (first entry)

XX DE Human liver peptide, SEQ ID No 37837.

KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.

XX OS Homo sapiens.

XX PN WO200157273-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000664.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488898/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 XX gene expression in human adult liver.

XX PS Claim 27; SEQ ID No 37837; 658pp; English.

XX CC The invention relates to a single exon nucleic acid probe (SENP) (I) for
 XX measuring human gene expression in a sample derived from human adult
 XX liver, comprising one of 13109 defined nucleotide sequences given in the
 XX specification (or complements/ fragments). The probe hybridises at high
 XX stringency to a nucleic acid molecule expressed in the human adult liver.
 XX (I) may be used for predicting, measuring and displaying gene expression
 XX in samples derived from human adult liver. The genes identified may be
 XX involved in genetic liver diseases such as cirrhosis,
 XX hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 XX associated with coronary heart disease. ABG47348-ABG59930 represent human
 XX liver single exon encoded peptides of the invention. Note: The sequence
 XX information for this patent does not appear in the printed specification
 XX but was obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 151 AA;

Query Match 100.0%; Score 37; DB 4; Length 151;
 Best Local Similarity 100.0%; Pred. No. 59;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 128 HHPHG 132

RESULT 19
 ABG46564
 ID ABG46564 standard; peptide; 151 AA.

XX AC ABG46564;

XX DT 19-AUG-2002 (first entry)

XX DE Human peptide encoded by genome-derived single exon probe SEQ ID 36229.

KW Human; single exon probe; asthma; lung cancer; COPD; ILD;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberosus sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
 KW primary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.

XX OS Homo sapiens.

XX PN WO200186003-A2.

XX PD 15-NOV-2001.

XX PF 30-JAN-2001; 2001WO-US000665.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2002-114183/15.

XX PT Spatially-addressable set of single exon nucleic acid probes, used to
 XX measure gene expression in human lung samples.

XX PS Claim 27; SEQ ID No 36229; 634pp; English.

XX CC The invention relates to a spatially-addressable set of single exon
 XX nucleic acid probes for measuring gene expression in a sample derived
 XX from human lung comprising single exon nucleic acid probes having one of
 XX 12614 nucleic acid sequences mentioned in the specification, or their
 XX complements or the 12387 open reading frames derived from the 12614
 XX probes. Also included are a microarray comprising the novel set of probes
 XX; the novel set of probes which hybridise at high stringency to a nucleic
 XX acid expressed in the human lung; measuring gene expression in a sample
 XX derived from human lung, comprising (a) contacting the array with a
 XX collection of detectably labeled nucleic acids derived from human lung
 XX mRNA, and (b) measuring the label detectably bound to each probe of the
 XX array; identifying exons in a eukaryotic genome, comprising (a)
 XX algorithmically predicting at least one exon from genomic sequences of
 XX the eukaryote; and (b) detecting specific hybridisation of detectably
 XX labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 XX having a fragment identical to the predicted exon, the probe is included
 XX in the above mentioned microarray; assigning exons to a single gene.

comprising (a) identifying exons from genomic sequence by the method above and (b) measuring the expression of each of the exons in several tissues and/or cell types using hybridisation to a single exon microarrays having a probe with the exon, where a common pattern of expression of the exons in the tissues and/or cell types indicates that the exons should be assigned to a single gene; a peptide comprising one of 12011 sequences, mentioned in the specification, or encoded by the probes/open reading frames (ORF). The probes are used for gene expression analysis, and for identifying exons in a gene, particularly using human lung derived mRNA and for the study of lung diseases such as asthma, lung cancer, chronic obstructive pulmonary disease (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension and hyaline membrane disease. The present sequence is a peptide/protein encoded by a single exon probe of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 100.0%; Score 37; DB 5; Length 151;
Best Local Similarity 100.0%; Pred. No. 59;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 128 HHPHG 132
|||||

RESULT 20

ABB71093
ID ABB71093 standard; protein; 198 AA.

AC ABB71093;

DT 26-MAR-2002 (first entry)

DE Drosophila melanogaster polypeptide SEQ ID NO 40071.

KW Drosophila; developmental biology; cell signalling; insecticide; pharmaceutical.

OS Drosophila melanogaster.

PN WO200171042-A2.

PD 27-SEP-2001.

PF 23-MAR-2001; 2001WO-US009231.

PR 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

PA (PEKE) PE CORP NY.

PI Venter JC, Adams M, Li PWD, Myers EW;

DR WPI; 2001-656860/75.

DR N-PSDB; ABL15196.

PT New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signaling and cell-cell interactions.

PS Disclosure; SEQ ID NO 40071; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and

cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 198 AA;

Query Match 100.0%; Score 37; DB 4; Length 198;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 153 HHPHG 157
|||||

RESULT 21

AAU55932
ID AAU55932 standard; protein; 248 AA.

AC AAU55932;

DT 27-FEB-2002 (first entry)

DE Propionibacterium acnes immunogenic protein #16828.

KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis; uveitis; endophthalmitis; bone; joint; central nervous system; ELISA; inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay; dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

PN WO200181581-A2.

PD 01-NOV-2001.

PF 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

PA (CORI-) CORIXA CORP.

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

DR WPI; 2001-616774/71.

DR N-PSDB; AAS59573.

PT Propionibacterium acnes polypeptides and nucleic acids useful for vaccinating against and diagnosing infections, especially useful for treating acne vulgaris.

PS Example 1; SEQ ID NO 17127; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic polypeptides. The proteins and their associated DNA sequences are used in the treatment, prevention and diagnosis of medical conditions caused by P. acnes. The disorders include SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in infections of bone, joints and the central nervous system, however it is particularly involved in the inflammatory lesions associated with acne vulgaris. A method for detecting the presence or absence of P. acnes in a patient comprises contacting a sample with a binding agent that binds to the proteins of the invention and determining the amount of bound protein in the sample. The polypeptides may be used as antigens in the production of antibodies specific for P. acnes proteins. These antibodies can be used to

CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 248 AA;

Query Match 100.0%; Score 37; DB 4; Length 248;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 142 HHPHG 146

RESULT 22
 ABMS2451
 ID ABMS2451 standard; protein; 248 AA.
 XX
 AC ABMS2451;
 XX
 DT 20-OCT-2003 (first entry)
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #17127.
 XX
 KW Acne vulgaris; antisborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 PA (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieve-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64502.
 XX
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 17127; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the

CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 248 AA;

Query Match 100.0%; Score 37; DB 6; Length 248;
 Best Local Similarity 100.0%; Pred. No. 98;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 142 HHPHG 146

RESULT 23
 ABU35012
 ID ABU35012 standard; protein; 259 AA.
 XX
 AC ABU35012;
 XX
 DT 19-JUN-2003 (first entry)
 DE Protein encoded by Prokaryotic essential gene #20539.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Moraxella catarrhalis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Travick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA38882.

XX New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 62936; 1766pp; English.
 XX The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for

CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

SQ Sequence 259 AA;
Query Match 100.0%; Score 37; DB 6; Length 259;
Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 HHPHG 5
Db 229 HHPHG 233

RESULT 24

ADL05482
ID ADL05482 standard; protein; 264 AA.

AC ADL05482;
XX
DT 06-MAY-2004 (first entry)
XX
DE M. catarrhalis protein #1248.
XX
KW Moraxella catarrhalis; infection.

XX Moraxella catarrhalis.

PN US6673910-B1.

PD 06-JAN-2004.

XX 04-APR-2000; 2000US-00540236.

PF 08-APR-1999; 99US-0128416P.

XX (GENO-) GENOME THERAPEUTICS CORP.

XX Breton GL;

XX WPI: 2004-178127/17.

DR N-PSDB; ADL03562.

PT New nucleic acid encoding a Moraxella catarrhalis polypeptide, useful for
PT preparing a composition for diagnosing, preventing or treating infection
PT caused by Moraxella catarrhalis.

PS Disclosure; SEQ ID NO 3168; 429pp; English.

XX The invention relates to an isolated nucleic acid encoding an Moraxella
CC catarrhalis polypeptide. The nucleic acid is useful for preparing a
CC composition for diagnosing, preventing or treating infection caused by
CC Moraxella catarrhalis. The present sequence represents the amino acid
CC sequence of a M. catarrhalis protein.

XX SQ Sequence 264 AA;
Query Match 100.0%; Score 37; DB 8; Length 264;
Best Local Similarity 100.0%; Pred. No. 1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 HHPHG 5
Db 230 HHPHG 234

RESULT 25

AAV32079
ID AAV32079 standard; protein; 273 AA.

XX AAV32079;

DT 17-JAN-2000 (first entry)

XX Human gap junction protein beta-4.

XX Gap junction protein beta-4; human; neurological disease;

KW epidermal disease; deafness; cataract; AIDS; therapy; diagnosis; vaccine.

XX Homo sapiens.

XX WO9951738-A1.

PD 14-OCT-1999.

XX 03-APR-1998; 98WO-CN000055.

PR 03-APR-1998; 98WO-CN000055.

XX (UYHU-) UNIV HUNAN MEDICAL.

XX Xia J, Pan Q, Liu C, Zheng D, Xie W;

DR WPI; 1999-611041/52.

DR N-PSDB; AA220259.

PT Novel polypeptides and polynucleotides useful for treating and preventing
PT e.g. neurological or epidermal diseases.

XX Claim 1; Page 25; 33pp; English.

XX The present sequence represents human gap junction protein beta-4, a
CC protein that has homology to connexins. The invention relates to gap
CC junction protein beta-4 polypeptides and polynucleotides (see also
CC AA220259), recombinant materials and methods for their production. It
CC also relates to methods for using such polypeptides and polynucleotides
CC for the treatment of neurological diseases, epidermal diseases, deafness,
CC cataract, and AIDS, and for identifying agonists, antagonists and/or
CC inhibitors useful in treating conditions associated with gap junction
CC protein beta-4 imbalance. Also provided are diagnostic assays for
CC detecting diseases associated with inappropriate gap junction protein
CC beta-4 activity or levels, e.g. by analysing for the presence or amount
CC of polypeptide expression in a sample. The polypeptides are also useful
CC for raising antibodies immunospecific to gap junction protein beta-4

XX SQ Sequence 273 AA;

Query Match 100.0%; Score 37; DB 2; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0

QY 1 HHPHG 5
Db 229 HHPHG 233

RESULT 26

AD117057
ID AD117057 standard; protein; 273 AA.
XX AC AD117057;
XX AC AD117057;
DT 15-APR-2004 (first entry)
XX
DE Human NOVX protein homologue SeqID 593.
XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (IgA) nephropathy; cirrhosis; arthritis;
KW Alzheimer's disease; infection; str.
XX
OS Homo sapiens.
XX
XX WO200268649-A2.
PN
XX
PD 06-SEP-2002.
XX
XX 31-JAN-2002; 2002WO-US0002785.
XX
PR 31-JAN-2001; 2001US-0265395P.
PR 31-JAN-2001; 2001US-0265412P.
PR 31-JAN-2001; 2001US-0265514P.
PR 31-JAN-2001; 2001US-0265517P.
PR 02-FEB-2001; 2001US-0266406P.
PR 05-FEB-2001; 2001US-0266767P.
PR 07-FEB-2001; 2001US-0266975P.
PR 08-FEB-2001; 2001US-0267057P.
PR 09-FEB-2001; 2001US-0267459P.
PR 15-FEB-2001; 2001US-0267823P.
PR 15-FEB-2001; 2001US-0268374P.
PR 26-FEB-2001; 2001US-0271664P.
PR 27-FEB-2001; 2001US-0271839P.
PR 27-FEB-2001; 2001US-0271855P.
PR 02-MAR-2001; 2001US-0272788P.
PR 02-MAR-2001; 2001US-0273046P.
PR 14-MAR-2001; 2001US-0275925P.
PR 14-MAR-2001; 2001US-0275947P.
PR 14-MAR-2001; 2001US-0275950P.
PR 14-MAR-2001; 2001US-0275989P.
PR 15-MAR-2001; 2001US-0276448P.
PR 15-MAR-2001; 2001US-0276450P.
PR 16-MAR-2001; 2001US-0276397P.
PR 16-MAR-2001; 2001US-0276768P.
PR 20-MAR-2001; 2001US-0278652P.
PR 26-MAR-2001; 2001US-0278775P.
PR 26-MAR-2001; 2001US-0278778P.
PR 29-MAR-2001; 2001US-0279882P.
PR 29-MAR-2001; 2001US-0279884P.
PR 30-MAR-2001; 2001US-0280147P.
PR 11-APR-2001; 2001US-0282992P.
PR 11-APR-2001; 2001US-0283083P.
PR 20-APR-2001; 2001US-0285133P.
PR 23-APR-2001; 2001US-0285749P.
PR 03-MAY-2001; 2001US-0288327P.
PR 03-MAY-2001; 2001US-0288504P.
PR 29-MAY-2001; 2001US-0294047P.
PR 30-MAY-2001; 2001US-0294473P.
PR 08-JUN-2001; 2001US-0296964P.
PR 18-JUN-2001; 2001US-0298959P.
PR 19-JUN-2001; 2001US-0299324P.
PR 13-AUG-2001; 2001US-0312020P.
PR 16-AUG-2001; 2001US-0312889P.
PR 16-AUG-2001; 2001US-0312908P.
PR 21-AUG-2001; 2001US-0313390P.
PR 28-AUG-2001; 2001US-0315470P.
PR 31-AUG-2001; 2001US-0316447P.
PR 07-SEP-2001; 2001US-0318115P.
PR 07-SEP-2001; 2001US-0318118P.
PR 12-SEP-2001; 2001US-0318740P.

PR 19-SEP-2001; 2001US-0323379P.
PR 18-OCT-2001; 2001US-0330245P.
PR 18-OCT-2001; 2001US-0330308P.
PR 14-NOV-2001; 2001US-0332701P.
XX
PA (CURA-) CURAGEN CORP.
XX
XX Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;
PI Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;
PI Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CFA;
PI Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;
XX WPI; 2002-706998/76.
DR
XX New NOVX polypeptides and nucleic acids, useful for preventing or
PT treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
PT atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
PT pharmacogenomics.
XX
PS Disclosure; SEQ ID NO 593; 1498pp; English.
XX
CC This invention relates to a novel nucleic acids, and encoded polypeptides
CC thereof, which have properties related to the stimulation of biochemical
CC or physiological responses in a cell, tissue, organ or organism.
CC Specifically, it refers to the use of biologically active fragments for
CC diagnostic and prognostic assays and furthermore in the treatment of
CC diverse pathological conditions. The present invention describes novel
CC human and murine NOVX proteins, as well as methods to modulate their
CC expression using antisense oligos, ribozymes and peptide nucleic acids.
CC The NOVX polypeptides, polynucleotides and antibodies are useful in
CC treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
CC atherosclerosis, cancer and diabetes. Furthermore, they may be used in
CC treating or preventing diseases such as inflammation, autoimmune
CC disorders, allergies, blood disorders, acquired immunodeficiency syndrome
CC (AIDS), obesity, asthma, immunoglobulin (IgA) nephropathy, cirrhosis,
CC arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
CC and epilepsy. Accordingly, these molecules have many activities including
CC haemostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
CC antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
CC neuroprotective, nootropic, antibacterial, virucide, antiparasitic,
CC relaxant and anticonvulsant. In addition, they are useful in screening
CC assays to identify small molecules that modulate or inhibit, for example,
CC neurogenesis, wound healing and angiogenesis. The nucleic acids are also
CC used as in chromosome mapping, tissue typing, preventive medicine and
CC pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
CC of the invention.
XX
SQ Sequence 273 AA;
Query Match 100.0%; Score 37; DB 5; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 229 HHPHG 233
|||||
RESULT 27
AD117060
ID AD117060 standard; protein; 273 AA.
XX
XX AC AD117060;
XX
DT 15-APR-2004 (first entry)
XX
DE Human NOVX protein homologue SeqID 596.
XX
XX human; NOVX; cardiomyopathy; atherosclerosis; cancer; diabetes;
KW inflammation; autoimmune disorder; allergy; blood disorder;
KW acquired immunodeficiency syndrome; AIDS; obesity; asthma;
KW immunoglobulin (IgA) nephropathy; cirrhosis; arthritis;
KW

KW	Alzheimer's disease; infection; str.	XX	XX	New NOVX polypeptides and nucleic acids, useful for preventing or
XX	Homo sapiens.	PT	PT	treating NOVX-associated disorders, e.g. cancer, cardiomyopathy,
XX	WO200268649-A2.	PT	PT	atherosclerosis, or diabetes, and in chromosome mapping, tissue typing or
XX		PT	PT	pharmacogenomics.
XX	06-SEP-2002.	XX	XX	
XX	31-JAN-2002; 2002WO-US002785.	XX	XX	Disclosure; SEQ ID NO 596; 1498pp; English.
XX	31-JAN-2001; 2001US-0265335P.	CC	CC	This invention relates to a novel nucleic acids, and encoded polypeptides
XX	31-JAN-2001; 2001US-0265412P.	CC	CC	thereof, which have properties related to the stimulation of biochemical
XX	31-JAN-2001; 2001US-0265514P.	CC	CC	or physiological responses in a cell, tissue, organ or organism.
XX	31-JAN-2001; 2001US-0265517P.	CC	CC	Specifically, it refers to the use of biologically active fragments for
XX	02-FEB-2001; 2001US-0266406P.	CC	CC	diagnostic and prognostic assays and furthermore in the treatment of
XX	03-FEB-2001; 2001US-0266767P.	CC	CC	diverse pathological conditions. The present invention describes novel
XX	07-FEB-2001; 2001US-0266975P.	CC	CC	human and murine NOVX proteins, as well as methods to modulate their
XX	08-FEB-2001; 2001US-0267459P.	CC	CC	expression using antisense oligos, ribozymes and peptide nucleic acids.
XX	09-FEB-2001; 2001US-0267823P.	CC	CC	The NOVX polypeptides, polynucleotides and antibodies are useful in
XX	15-FEB-2001; 2001US-0268974P.	CC	CC	treating or preventing NOVX-associated disorders, e.g. cardiomyopathy,
XX	26-FEB-2001; 2001US-0271664P.	CC	CC	atherosclerosis, cancer and diabetes. Furthermore, they may be used in
XX	27-FEB-2001; 2001US-0271839P.	CC	CC	treating or preventing diseases such as inflammation, autoimmune
XX	02-MAR-2001; 2001US-0272788P.	CC	CC	disorders, allergies, asthma, blood disorders, acquired immunodeficiency syndrome
XX	14-MAR-2001; 2001US-0275925P.	CC	CC	(AIDS), obesity, asthma, immunoglobulin (Ig)A nephropathy, cirrhosis,
XX	14-MAR-2001; 2001US-0275947P.	CC	CC	arthritis, Alzheimer's disease, infections, stroke, muscular dystrophy
XX	14-MAR-2001; 2001US-0275950P.	CC	CC	and epilepsy. Accordingly, these molecules have many activities including
XX	15-MAR-2001; 2001US-0276448P.	CC	CC	cytostatic, cardiant, antiinflammatory, immunosuppressive, antiallergic,
XX	15-MAR-2001; 2001US-0276450P.	CC	CC	haemostatic, anti-HIV, antidiabetic, antiarthritic, hepatotropic,
XX	16-MAR-2001; 2001US-0276397P.	CC	CC	antiasthmatic, nephrotropic, antiarthritic, hepatotropic,
XX	20-MAR-2001; 2001US-0276768P.	CC	CC	neuroprotective, nontropic, antibacterial, virucide, antiparasitic,
XX	26-MAR-2001; 2001US-0278652P.	CC	CC	relaxant and anticonvulsant. In addition, they are useful in screening
XX	26-MAR-2001; 2001US-0278778P.	CC	CC	assays to identify small molecules that modulate or inhibit, for example,
XX	29-MAR-2001; 2001US-0279882P.	CC	CC	neurogenesis, wound healing and angiogenesis. The nucleic acids are also
XX	29-MAR-2001; 2001US-0279884P.	CC	CC	used as in chromosome mapping, tissue typing, preventive medicine and
XX	11-APR-2001; 2001US-0280147P.	CC	CC	pharmacogenomics. This polypeptide is a homologue of a human NOVX protein
XX	11-APR-2001; 2001US-0282992P.	CC	CC	of the invention.
XX	20-APR-2001; 2001US-0283083P.	XX	XX	
XX	23-APR-2001; 2001US-0285133P.	XX	XX	Sequence 273 AA;
XX	03-MAY-2001; 2001US-0285749P.	QY	QY	Query Match 100.0%; Score 37; DB 5; Length 273;
XX	29-MAY-2001; 2001US-0288504P.	Db	Db	Best Local Similarity 100.0%; Pred. No. 1.1e+02;
XX	30-MAY-2001; 2001US-0294047P.			Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX	08-JUN-2001; 2001US-0296864P.			
XX	18-JUN-2001; 2001US-0298959P.			
XX	19-JUN-2001; 2001US-0299324P.			
XX	13-AUG-2001; 2001US-0312020P.			
XX	16-AUG-2001; 2001US-0312889P.			
XX	16-AUG-2001; 2001US-0312908P.			
XX	21-AUG-2001; 2001US-0313390P.			
XX	28-AUG-2001; 2001US-0315470P.			
XX	31-AUG-2001; 2001US-0316447P.			
XX	07-SEP-2001; 2001US-0318115P.			
XX	12-SEP-2001; 2001US-0318118P.			
XX	19-SEP-2001; 2001US-0318740P.			
XX	18-OCT-2001; 2001US-0323379P.			
XX	18-OCT-2001; 2001US-0330245P.			
XX	14-NOV-2001; 2001US-0332701P.			
XX	(CURA-) CURAGEN CORP.			
XX	Tchernev VT, Spytek KA, Zerhusen BD, Patturajan M, Shimkets RA;			
XX	Li L, Gangolli EA, Padigar M, Anderson DW, Rastelli L, Miller CE;			
XX	Gerlach VL, Taupier RJ, Gusev VY, Colman SD, Wolenc AR, Pena CE;			
XX	Furtak K, Grosse WM, Alsobrook JP, Lepley DM, Rieger DK, Burgess CE;			
XX	WPI; 2002-706998/76.			

XX Mack DH, Aziz N;
 XX WPI: 2003-201532/19.
 DR N-PSDB; ACC51048.
 DR
 XX Detecting a bladder cancer-associated transcript in a cell from a
 PT patient, comprises contacting a biological sample from the patient with a
 PT bladder cancer-associated polynucleotide or antibody.
 XX
 XX Claim 10; Page 293; 307pp; English.
 PS
 XX The present invention describes a method for detecting a bladder cancer-
 CC associated transcript in a cell from a patient. The method comprises
 CC contacting a biological sample from the patient with a polynucleotide
 CC that selectively hybridizes to a sequence that is 80 % identical to a
 CC table of sequences (see ACC50951 to ACC51059). ACC50951 to ACC51059
 CC encode the human bladder cancer-associated proteins given in ABR48146 to
 CC ABR48242). Bladder cancer-associated sequences from the present invention
 CC have cytostatic activities, and can be used in antisense gene therapy and
 CC in vaccine production. The method can be used for detecting a bladder
 CC cancer-associated transcript in a cell from a patient. The method is
 CC useful in diagnosing or treating bladder cancer and in screening for
 CC compounds that modulate bladder cancer, such as hormones or antibodies.
 CC The nucleic acid molecules from the present invention may be used in
 CC various screening and diagnostic methods, and for gene therapy, vaccine
 CC and/or antisense/inhibition applications
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 100.0%; Score 37; DB 6; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 DB 229 HHPHG 233
 RESULT 29
 ABUS6449
 ID ABUS6449 standard; protein; 273 AA.
 AC ABUS6449;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #42.
 XX
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 XX WO200286443-A2.
 PN
 XX
 XX 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 29-NOV-2001; 2001US-0334370P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PR
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Aziz N, Murray R;
 PI

XX WPI: 2003-093161/08.
 DR N-PSDB; ABX76170.
 XX
 PT Detecting a lung cancer-associated transcript in a cell from a patient
 PT for treating lung cancer, by contacting a biological sample from the
 PT patient with a polynucleotide that exhibits increased or decreased
 PT expression in lung cancer.
 XX
 XX Claim 27; Page 222; 453pp; English.
 PS
 XX The invention relates to a method for detecting a lung cancer-associated
 CC transcript in a cell from a patient, comprising contacting a biological
 CC sample from the patient with a polynucleotide that selectively hybridizes
 CC to a sequence that is at least 80 % identical to a gene that exhibits
 CC increased or decreased expression in lung cancer samples. Lung cancer-
 CC associated polynucleotides and polypeptides are used for identifying a
 CC compound that modulates a lung cancer-associated polypeptide, for
 CC inhibiting proliferation of a lung cancer-associated cell to treat lung
 CC cancer in a patient and for treating a mammal having lung cancer by
 CC administering a modulatory compound identified. The methods are useful
 CC for treating lung cancer, such as small cell lung cancer, non-small cell
 CC lung cancer or other benign or precancerous lesions, e.g. atelectasis,
 CC emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis,
 CC hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and
 CC bronchiectasis. The genes, polynucleotides and polypeptides are useful
 CC for diagnostic purposes and as targets for screening for therapeutic
 CC compounds that modulate lung cancer, such as antibodies. Sequences
 CC ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the
 CC invention
 XX
 XX Sequence 273 AA;
 SQ
 Query Match 100.0%; Score 37; DB 6; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 DB 229 HHPHG 233
 RESULT 30
 ABUS6673
 ID ABUS6673 standard; protein; 273 AA.
 AC ABUS6673;
 XX
 DT 02-APR-2003 (first entry)
 XX
 DE Lung cancer-associated polypeptide #266.
 XX
 KW Lung cancer-associated polypeptide; cytostatic; emphysema;
 KW antiinflammatory; antiasthmatic; non-small cell lung cancer; atelectasis;
 KW small cell lung cancer; benign lesion; precancerous lesion; bronchitis;
 KW chronic obstructive pulmonary disease; hypersensitivity pneumonitis;
 KW interstitial pulmonary fibrosis; fibrosis; asthma; bronchiectasis.
 XX
 OS Unidentified.
 XX
 XX WO200286443-A2.
 PN
 XX
 XX 31-OCT-2002.
 XX
 XX 18-APR-2002; 2002WO-US012476.
 PF
 XX 18-APR-2001; 2001US-0284770P.
 PR
 XX 10-MAY-2001; 2001US-0290492P.
 PR
 XX 09-NOV-2001; 2001US-0339245P.
 PR
 XX 13-NOV-2001; 2001US-0350666P.
 PR
 XX 29-NOV-2001; 2001US-0334370P.
 PR
 XX 12-APR-2002; 2002US-0372246P.
 PR
 XX (EOSB-) EOS BIOTECHNOLOGY INC.
 PA
 XX Aziz N, Murray R;
 PI

(EOSB-) EOS BIOTECHNOLOGY INC.
 Aziz N, Murray R;
 WPI; 2003-093161/08.
 N-PSDB; ABX76402.
 Detecting a lung cancer-associated transcript in a cell from a patient for treating lung cancer, by contacting a biological sample from the patient with a polynucleotide that exhibits increased or decreased expression in lung cancer.
 Claim 27; Page 393; 453pp; English.
 The invention relates to a method for detecting a lung cancer-associated transcript in a cell from a patient, comprising contacting a biological sample from the patient with a polynucleotide that selectively hybridises to a sequence that is at least 80 % identical to a gene that exhibits increased or decreased expression in lung cancer samples. Lung cancer-associated polynucleotides and polypeptides are used for identifying a compound that modulates a lung cancer-associated polypeptide, for inhibiting proliferation of a lung cancer-associated cell to treat lung cancer in a patient and for treating a mammal having lung cancer by administering a modulatory compound identified. The methods are useful for treating lung cancer, such as small cell lung cancer, non-small cell lung cancer or other benign or precancerous lesions, e.g. atelectasis, emphysema, bronchitis, chronic obstructive pulmonary disease, fibrosis, hypersensitivity pneumonitis, interstitial pulmonary fibrosis, asthma and bronchiectasis. The genes, polynucleotides and polypeptides are useful for diagnostic purposes and as targets for screening for therapeutic compounds that modulate lung cancer, such as antibodies. Sequences ABUS6408-ABUS6745 represent lung cancer-associated polypeptides of the invention
 Sequence 273 AA;
 Query Match 100.0%; Score 37; DB 6; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 |||||
 Db 229 HHPHG 233
 RESULT 31
 ADN39026
 ID ADN39026 standard; protein; 273 AA.
 AC ADN39026;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:344.
 DE
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW retinal neovascularisation syndrome; scarring; uterine fibroid;
 KW detection; diagnosis; prognosis; drug screening; drug targeting;
 KW wound healing; contraception; cytostatic; cardiac; immunomodulatory;
 KW vulnary; gene therapy; vaccine.
 XX
 OS Homo sapiens.
 XX
 PN WO2003042661-A2.
 XX
 PD 22-MAY-2003.
 XX
 PF 13-NOV-2002; 2002WO-US036810.
 XX
 PR 13-NOV-2001; 2001US-0350666P.
 PR 21-NOV-2001; 2001US-0332464P.
 PR

29-NOV-2001; 2001US-0334393P.
 03-DEC-2001; 2001US-0335394P.
 14-DEC-2001; 2001US-0340376P.
 08-JAN-2002; 2002US-0347211P.
 10-JAN-2002; 2002US-0347349P.
 08-FEB-2002; 2002US-0355250P.
 13-FEB-2002; 2002US-0356714P.
 20-FEB-2002; 2002US-0359077P.
 29-MAR-2002; 2002US-0368809P.
 04-APR-2002; 2002US-0370110P.
 12-APR-2002; 2002US-0372246P.
 05-JUN-2002; 2002US-0386614P.
 16-JUL-2002; 2002US-0396839P.
 22-JUL-2002; 2002US-0397775P.
 22-JUL-2002; 2002US-0397845P.
 09-SEP-2002; 2002US-0409450P.
 (EOSB-) EOS BIOTECHNOLOGY INC.
 Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;
 Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;
 WPI; 2003-468649/44.
 N-PSDB; ADN39025.
 Determining the presence or absence of a pathological cell in a patient, useful for diagnosing, prognosing or treating cancer, comprises detecting a nucleic acid in a biological sample.
 Claim 12; SEQ ID NO 344; 1385pp; English.
 The invention relates to nucleic acids and proteins (ADN38683-ADN40064) whose expression is upregulated or downregulated in specific cancers or other diseases such as angiogenic or fibrotic disorders, and to methods of determining the presence or absence of a pathological cell in a patient by detecting a nucleic acid at least 80% identical to those of the invention or by detecting a polypeptide of the invention. The invention also relates to expression vectors and host cells comprising a nucleic acid of the invention; antibodies which specifically bind a polypeptide of the invention; use of such antibodies for drug targeting; and methods of screening for modulators of activity or expression of the polypeptides and nucleic acids. The nucleic acids, polypeptides, antibodies and methods are useful for diagnosing, prognosing and treating cancer and other conditions such as psoriasis, ischaemia, heart disease, atherosclerosis, inflammatory diseases, autoimmune diseases, retinal neovascularisation syndromes, scarring and uterine fibroids. They may also be useful in wound healing and in contraception. The present sequence represents a polypeptide of the invention.
 Sequence 273 AA;
 Query Match 100.0%; Score 37; DB 7; Length 273;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 |||||
 Db 229 HHPHG 233
 RESULT 32
 ADN39932
 ID ADN39932 standard; protein; 273 AA.
 AC ADN39932;
 XX
 DT 17-JUN-2004 (first entry)
 XX
 DE Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:C302.
 DE
 KW Human; differential expression; cancer; angiogenic disorder;
 KW fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;
 KW inflammatory disease; autoimmune disease;
 KW

KW retinal neovascularisation syndrome; scarring; uterine fibroid;
KW detection; diagnosis; prognosis; drug screening; drug targeting;
KW wound healing; contraception; cytostatic; cardiant; immunomodulatory;
KW vulnery; gene therapy; vaccine.

OS Homo sapiens.

PN WO2003042661-A2.

XX 22-MAY-2003.

XX 13-NOV-2002; 2002WO-US036810.

XX 13-NOV-2001; 2001US-0350666P.

PR 21-NOV-2001; 2001US-0332464P.

PR 29-NOV-2001; 2001US-0334393P.

PR 03-DEC-2001; 2001US-0335394P.

PR 14-DEC-2001; 2001US-0340376P.

PR 08-JAN-2002; 2002US-0347211P.

PR 10-JAN-2002; 2002US-0347349P.

PR 08-FEB-2002; 2002US-0355250P.

PR 13-FEB-2002; 2002US-0356714P.

PR 20-FEB-2002; 2002US-0359077P.

PR 29-MAR-2002; 2002US-0368609P.

PR 04-APR-2002; 2002US-0370110P.

PR 12-APR-2002; 2002US-0372246P.

PR 05-JUN-2002; 2002US-0386614P.

PR 16-JUL-2002; 2002US-0396839P.

PR 22-JUL-2002; 2002US-0397775P.

PR 22-JUL-2002; 2002US-0397845P.

PR 09-SEP-2002; 2002US-0409450P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Ariz N, Ginsburg WM, Gish KC, Glynne R, Heverzi PA;

PI Mack DH, Murray R, Watson SR, Wilson KE, Zlotnik A;

XX WPI; 2003-468649/44.

DR N-PSDB; ADN39715.

XX Determining the presence or absence of a pathological cell in a patient,
PT useful for diagnosing, prognosing or treating cancer, comprises detecting
PT a nucleic acid in a biological sample.

XX Claim 12; SEQ ID NO C302; 1385pp; English.

PS The invention relates to nucleic acids and proteins (ADN38683-ADN40064)
CC whose expression is upregulated or downregulated in specific cancers or
CC other diseases such as angiogenic or fibrotic disorders, and to methods
CC of determining the presence or absence of a pathological cell in a
CC patient by detecting a nucleic acid at least 80% identical to those of
CC the invention or by detecting a polypeptide of the invention. The
CC invention also relates to expression vectors and host cells comprising a
CC nucleic acid of the invention; antibodies which specifically bind a
CC polypeptide of the invention; use of such antibodies for drug targeting;
CC and methods of screening for modulators of activity or expression of the
CC polypeptides and nucleic acids. The nucleic acids, polypeptides,
CC antibodies and methods are useful for diagnosing, prognosing and treating
CC cancer and other conditions such as psoriasis, ischaemia, heart disease,
CC atherosclerosis, inflammatory diseases, autoimmune diseases, retinal
CC neovascularisation syndromes, scarring and uterine fibroids. They may
CC also be useful in wound healing and in contraception. The present
CC sequence represents a polypeptide of the invention.

XX Sequence 273 AA;

Query Match 100.0%; Score 37; DB 7; Length 273;
Best Local Similarity 100.0%; Pred. NO. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 229 HHPHG 233

RESULT 33

ABB69468

ID ABB69468 standard; protein; 323 AA.

XX AC ABB69468;

XX 26-MAR-2002 (first entry)

XX Drosophila melanogaster polypeptide SEQ ID NO 35196.

KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical.

OS Drosophila melanogaster.

XX WO200171042-A2.

XX 27-SEP-2001.

XX 23-MAR-2001; 2001WO-US009231.

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

XX (PEKE) PE CORP NY.

XX Venter JC, Adams M, Li PWD, Myers EW;

XX WPI; 2001-656860/75.

DR N-PSDB; ABL13571.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT genes from Drosophila and for elucidating cell signaling and cell-cell

PT interactions.

XX Disclosure; SEQ ID NO 35196; 21pp + Sequence Listing; English.

XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 323 AA;

Query Match 100.0%; Score 37; DB 4; Length 323;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 279 HHPHG 283

RESULT 34

ADN46756

ID ADN46756 standard; protein; 326 AA.

XX AC ADN46756;

XX 01-JUL-2004 (first entry)

XX Thermococcus kodakaraensis KOD1 protein sequence SeqID634.

XX gene disruption; gene targeting; marker gene; transformation;
KW homologous recombination; hyperthermostable archaeobacterium; KOD1;

KW gene structure; gene function; enzyme activity; medicine;
 KW forensic science; food; drug inspection; molecular biology; immunology.
 XX Thermococcus kodakaraensis.
 OS WO2004022736-A1.
 PN 18-MAR-2004.
 PD
 XX 29-AUG-2003; 2003WO-IB003597.
 PF
 XX 30-AUG-2002; 2002JP-00319011.
 PR
 XX (NISC-) JAPAN SCI & TECHNOLOGY CORP.
 PA Imanaka T, Atomi H;
 PI WPI; 2004-257583/24.
 XX
 DR Method for disrupting targeted gene in genome of organism particularly
 PT thermostable bacterium and with genome chips for analysis, applicable in
 PT studying gene structure and functions.
 XX
 PS Claim 9; SEQ ID NO 634; 598pp; Japanese.
 XX
 CC This invention relates to a novel method for targeting disruption of an
 CC arbitrary gene in a genome of an organism which comprises providing the
 CC whole sequential data of the genome of such organism, selecting at least
 CC 1 arbitrary region in the sequence, providing a vector that contains a
 CC sequence homologous with the selected region and a marker gene,
 CC transformation, and homologous recombination. The genome is preferably
 CC the genome of a hyperthermostable archaeobacterium, particularly
 CC Thermococcus kodakaraensis KOD1. The method is for targeting the
 CC disruption of a gene in the genome of an organism, which is applicable in
 CC studying gene structure and functions as well as enzyme activities of
 CC encoded proteins and useful in medicine, forensic science, food or drug
 CC inspection, molecular biology and immunology. With this method, the
 CC disruption of a gene at an arbitrary position in a genome can be achieved
 CC efficiently and reliably. The present sequence is that of a protein
 CC encoded by the genome of Thermococcus kodakaraensis which was derived
 CC using the method of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 326 AA;
 Query Match 100.0%; Score 37; DB 8; Length 326;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 Db 138 HHPHG 142
 RESULT 35
 ABB66932
 ID ABB66932 standard; protein; 327 AA.
 AC ABB66932;
 XX
 XX 26-MAR-2002 (first entry)
 DT
 DE Drosophila melanogaster polypeptide SEQ ID NO 27588.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 PN
 XX

PD 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 PA Venter JC, Adams M, Li PWD, Myers EW;
 PI WPI; 2001-656860/75.
 DR N-PSDB; ABL11035.
 DR
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 27588; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 327 AA;
 Query Match 100.0%; Score 37; DB 4; Length 327;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HHPHG 5
 Db 143 HHPHG 147
 RESULT 36
 ADS30838
 ID ADS30838 standard; protein; 350 AA.
 XX
 XX ADS30838;
 XX
 DT 02-DEC-2004 (first entry)
 XX
 DE Bacterial polypeptide #19871.
 XX
 XX Recombinant DNA construct; improved plant property;
 KW cold tolerance; heat tolerance; drought tolerance; herbicide; osmosis;
 KW pathogen tolerance; pest tolerance; plant disease resistance;
 KW cell cycle pathway modification; plant growth regulator;
 KW homologous recombination; seed oil yield; protein yield; carbohydrate;
 KW nitrogen; phosphorus; photosynthesis; lignin; galactomannan;
 KW bacterial polypeptide.
 XX
 OS Bacteria.
 XX
 PN US2003233675-A1.
 XX
 PD 18-DEC-2003.
 XX
 XX 20-FEB-2003; 2003US-00369493.
 PF
 XX 21-FEB-2002; 2002US-0360039P.
 PR
 XX (CAOY/) CAO Y.
 PA (HINK/) HINKLE G J.
 PA (SLAT/) SLATER S C.

PA (CHEN/) CHEN X.
 PA (GOLD/) GOLDMAN B S.
 XX Cao Y, Hinkle GJ, Slater SC, Chen X, Goldman BS;
 XX WPI; 2004-061375/06.
 DR
 XX New recombinant DNA construct comprising a promoter positioned to provide
 PT for expression of a polynucleotide encoding a polypeptide from a
 PT microbial source, useful for producing plants with improved properties.
 XX
 XX Claim 1; SEQ ID NO 19871; 122pp; English.
 PS
 XX The invention relates to a recombinant DNA construct comprising a
 CC promoter functional in a plant cell, where the promoter is positioned to
 CC provide for expression of a polynucleotide encoding a polypeptide from a
 CC microbial source. The invention also relates to a transformed plant
 CC comprising the recombinant DNA construct and a method of producing a
 CC transformed plant having an improved property. The plant is a crop plant
 CC such as maize or soybean. The method of producing a transformed plant
 CC having an improved property comprises transforming a plant with the
 CC recombinant DNA construct and growing the transformed plant, where the
 CC polynucleotide or polypeptide is useful for improving plant properties.
 CC The recombinant DNA construct is useful for producing plants with
 CC improved plant properties, e.g. improved cold, heat or drought tolerance,
 CC tolerance to herbicides, extreme osmotic conditions, pathogens or pests,
 CC increased resistance to plant disease, better growth rate by modification
 CC of the cell cycle pathway with plant growth regulators, increased rate of
 CC homologous recombination, modified seed oil or protein yield and/or
 CC content, improved yield by modification of carbohydrate, nitrogen or
 CC phosphorus use and/or uptake, by modification of photosynthesis or by
 CC providing improved plant growth and development under at least one stress
 CC condition, improved lignin production or improved galactomannan
 CC production. This sequence represents a bacterial polypeptide used in the
 CC scope of the invention. Note: The sequence data for this patent did not
 CC form part of the printed specification but was obtained in electronic
 CC format from USPTO at seqdata.uspto.gov/sequence.html.
 XX
 XX Sequence 350 AA;
 SQ
 Query Match 100.0%; Score 37; DB 8; Length 350;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HHPHG 5
 Db 174 HHPHG 178
 |||||
 RESULT 37
 ABB71820
 ID ABB71820 standard; protein; 402 AA.
 XX
 AC ABB71820;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 42252.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US0009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL03582.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 42252; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABLI16176-ABLI30511), expressed DNA
 CC sequences (ABLI01840-ABLI16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 402 AA;
 Query Match 100.0%; Score 37; DB 4; Length 402;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 HHPHG 5
 Db 160 HHPHG 164
 |||||
 RESULT 38
 ABB59479
 ID ABB59479 standard; protein; 419 AA.
 XX
 AC ABB59479;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 5229.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US0009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL03582.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 5229; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
CC ABB72072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 419 AA;

Query Match 100.0%; Score 37; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 253 HHPHG 257

RESULT 39
ABO77799
ID ABO77799 standard; protein; 466 AA.
AC ABO77799;
XX
DT 29-JUL-2004 (first entry)
XX
DE Pseudomonas aeruginosa polypeptide #9974.
DE Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
KW Pseudomonas aeruginosa.
XX
OS US6551795-B1.
PN
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX

PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
DR WPI; 2003-615309/58.
DR N-PSDB; ABD11370.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 26545; 455pp; English.
XX

CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biotech technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html

XX Sequence 466 AA;

Query Match 100.0%; Score 37; DB 7; Length 466;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 135 HHPHG 139

RESULT 40
ABB79804
ID ABB79804 standard; protein; 525 AA.

XX ABB79804;
AC
XX
DT 25-NOV-2002 (first entry)
XX

DE Human histidine proline rich glycoprotein.
XX
KW Histidine proline rich glycoprotein; HPRG; human; antiangiogenic;
KW cystostatic; antiarteriosclerotic; antiinflammatory; antidiabetic;
KW gynaecological; antiarthritic; antiulcer; osteopathic; antitumour;
KW ophthalmological; nootropic; neuroprotective; antiparkinsonian; chelator.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Domain 276..321
FT Domain /note= "proline-rich domain"
FT Domain 350..497
FT Domain /note= "Histidine-proline-rich domain, region also
FT specifically claimed in Claim 1"

XX WO200264621-A2.

XX 22-AUG-2002.

XX 14-FEB-2002; 2002WO-US004336.

XX 14-FEB-2001; 2001US-0268370P.

XX (ATTE-) ATTENUON LLC.

XX Donate F, Harris S, Plunkett ML, Mazar AP;

XX WPI; 2002-666989/71.

XX P-PSDB; ABN84910.

XX New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX Claim 1; Page 11; 82pp; English.

XX The present sequence is the protein sequence of human histidine proline
CC rich glycoprotein (HPRG), a proteinaceous metal chelator that can be used
CC to inhibit angiogenesis and treat cancer. Claimed anti-angiogenic
CC polypeptides or peptides comprise: the histidine-proline-rich (H/P)
CC domain (see ABB79806) of human HPRG; an H/P domain of rabbit HPRG (see
CC ABB79807); a variant of these that is capable of inhibiting angiogenesis,
CC endothelial cell proliferation or endothelial tube formation in vitro or
CC in vivo; or a pentapeptide having the generic sequence given in ABB79808.
CC Also claimed are: chemically synthesised or recombinantly produced
CC peptide multimers; a diagnostically or therapeutically labeled anti-
CC angiogenic polypeptide, peptide or peptide multimer; a diagnostically
CC useful HPRG-related composition, comprising the diagnostically labeled
CC polypeptide, peptide or peptide multimer and a carrier; an antibody
CC specific for an epitope of HPRG that is present in the H/P domain of
CC human or rabbit HPRG, and which binds to HPRG or to any of the domains in
CC a way which inhibits the anti-angiogenic activity of HPRG or the domain,

CC or an antigen-binding fragment of the antibody; a method for inhibiting
CC cell migration, cell invasion, cell proliferation or angiogenesis, or for
CC inducing apoptosis; a method for treating a subject having a disease or
CC condition associated with undesired cell migration, invasion,
CC proliferation, or angiogenesis; a method of stimulating or inhibiting
CC angiogenesis in a subject; a method of detecting the presence of HPRG or
CC its cleavage product or its peptide in a biological sample; isolated
CC nucleic acids encoding the polypeptide, peptide or peptide multimer; an
CC expression vector; transformed or transfected cells; a method of
CC providing to a cell, tissue or organ an angiogenesis-inhibitory amount of
CC HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer;
CC an affinity ligand useful for binding to, or isolating, an HPRG-binding
CC molecule or cells expressing the binding molecule, comprising the
CC polypeptide, peptide or peptide multimer, immobilised to a solid support
CC or carrier; and a method of isolating HPRG-binding molecule, or isolating
CC or enriching cells expressing HPRG-binding site or receptor, from a
CC complex mixture. The compositions and methods are useful in diagnosing or
CC treating a disease or condition associated with undesired cell migration,
CC invasion, proliferation, or angiogenesis, such as cancer,
CC atherosclerosis, diabetic retinopathy, inflammation, endometriosis,
CC arthritis, peptic ulcers, or fractures. HPRG is especially useful in
CC inhibiting the growth of primary tumours or metastases, and may also be
CC used in treating neurodegenerative diseases like Alzheimer's disease,
CC Parkinson's disease, and amyotrophic lateral sclerosis. The antibodies
CC are stimulators of angiogenesis and are useful for promoting
CC neovascularization in pertinent disease states, and in various
CC immunoassays
XX
SQ

Sequence 525 AA;

Query Match 100.0%; Score 37; DB 5; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

Db 384 HHPHG 388

Search completed: June 15, 2005, 14:15:48
Job time : 93 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-8

Perfect score: 37

Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1 pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	6	3	US-09-095-407-3
2	37	100.0	11	3	US-09-095-407-1
3	37	100.0	11	4	US-09-742-494-1
4	37	100.0	11	4	US-09-809-657-1
5	37	100.0	26	3	US-08-785-636-1
6	37	100.0	26	3	US-09-095-407-2
7	37	100.0	26	4	US-09-742-494-2
8	37	100.0	26	4	US-09-809-657-2
9	37	100.0	69	4	US-09-248-796A-24357
10	37	100.0	163	4	US-09-302-540-13395
11	37	100.0	220	4	US-09-270-767-61056
12	37	100.0	264	4	US-09-540-236-3168
13	37	100.0	273	4	US-09-949-016-6606
14	37	100.0	280	4	US-09-949-016-9101
15	37	100.0	466	4	US-09-252-991A-26545
16	37	100.0	525	4	US-09-976-594-64
17	37	100.0	525	4	US-09-919-039-62
18	37	100.0	539	4	US-09-518-959-8
19	37	100.0	539	4	US-09-518-959-9
20	37	100.0	618	4	US-09-252-991A-23373
21	37	100.0	810	4	US-09-248-796A-20281
22	37	100.0	923	4	US-09-270-767-45546
23	31	83.8	5	4	US-09-546-013-29
24	31	83.8	7	3	US-08-776-265-12
25	31	83.8	7	4	US-09-398-184-12
26	31	83.8	9	3	US-08-159-339A-1005
27	31	83.8	15	3	US-09-082-092-14

28	31	83.8	15	4	US-09-885-722A-14	Sequence 14, Appl
29	31	83.8	52	3	US-09-227-357-575	Sequence 575, App
30	31	83.8	69	4	US-09-910-009A-158	Sequence 158, App
31	31	83.8	75	4	US-09-716-129-73	Sequence 73, Appl
32	31	83.8	76	4	US-09-489-039A-10688	Sequence 10688, A
33	31	83.8	84	4	US-09-270-767-57094	Sequence 57094, A
34	31	83.8	91	4	US-09-270-767-36032	Sequence 36032, A
35	31	83.8	91	4	US-09-270-767-51249	Sequence 51249, A
36	31	83.8	92	4	US-09-248-796A-23534	Sequence 23534, A
37	31	83.8	99	4	US-09-640-211A-653	Sequence 653, App
38	31	83.8	102	4	US-09-270-767-58449	Sequence 58449, A
39	31	83.8	115	4	US-09-461-325-160	Sequence 160, App
40	31	83.8	115	4	US-10-012-542-160	Sequence 160, App
41	31	83.8	115	4	US-10-115-123-160	Sequence 24231, A
42	31	83.8	125	4	US-09-248-796A-24231	Sequence 160, App
43	31	83.8	127	4	US-09-430-221-2	Sequence 2, Appli
44	31	83.8	129	4	US-09-205-258-1135	Sequence 1135, Ap
45	31	83.8	132	1	US-08-470-179-26	Sequence 26, Appl

ALIGNMENTS

RESULT 1
US-09-095-407-3
; Sequence 3, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; Yip, Tai-Tung
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSER: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-095-407-3

```
Query Match      100.0%; Score 37; DB 3; Length 6;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      2 HHPHG 6

RESULT 2
US-09-095-407-1
; Sequence 1, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 11 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-095-407-1

Query Match      100.0%; Score 37; DB 3; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      2 HHPHG 6

RESULT 3
US-09-742-494-1
; Sequence 1, Application US/09742494
; Patent No. 6528320
; GENERAL INFORMATION:

Query Match      100.0%; Score 37; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      2 HHPHG 6

RESULT 4
US-09-809-657-1
; Sequence 1, Application US/09809657
; Patent No. 6734022
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-1

Query Match      100.0%; Score 37; DB 4; Length 11;
Best Local Similarity 100.0%; Pred. No. 0.83;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      2 HHPHG 6

RESULT 5
US-08-785-636-1
; Sequence 1, Application US/08785636
; Patent No. 6027942
; GENERAL INFORMATION:
; APPLICANT: Yip, Tai-Tung
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Method and Apparatus for Desorption and Ionization of
; Analytes
; FILE REFERENCE: Hutchens
; CURRENT APPLICATION NUMBER: US/08/785,636
; CURRENT FILING DATE: 1997-01-17
; EARLIER APPLICATION NUMBER: 08/068,896
; EARLIER FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Unknown
```

```
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: peptide
US-08-785-636-1

Query Match      100.0%; Score 37; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 6
US-09-095-407-2
; Sequence 2, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-095-407-2

Query Match      100.0%; Score 37; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 7
US-09-095-407-2
; Sequence 2, Application US/09095407
; Patent No. 6124137
; GENERAL INFORMATION:
; APPLICANT: Hutchens, T. William
; TITLE OF INVENTION: Surface-Enhanced Neat Desorption for the
; Detection of Analytes
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fulbright & Jaworski L.L.P.
; STREET: 1301 McKinney, Suite 5100
; CITY: Houston
; STATE: Texas
; COUNTRY: USA
; ZIP: 77010-3095
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA: US/09/095,407
; FILING DATE: 10-Jun-1998
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/068,896
; FILING DATE: 28-MAY-1993
; APPLICATION NUMBER: WO PCT/US94/06064
; FILING DATE: 27-MAY-1994
; APPLICATION NUMBER: US 08/556,951
; FILING DATE: 27-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul, Thomas D.
; REGISTRATION NUMBER: 32,714
; REFERENCE/DOCKET NUMBER: D-5639-PCT-US-D1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (713) 651-5325
; TELEFAX: (713) 651-5246
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 26 amino acids
; TYPE: amino acid
; STRANDEDNESS: <Unknown>
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-095-407-2

Query Match      100.0%; Score 37; DB 3; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 8
US-09-809-657-2
; Sequence 2, Application US/09809657
; Patent No. 6734022
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; PRIOR FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-2

Query Match      100.0%; Score 37; DB 4; Length 26;
Best Local Similarity 100.0%; Pred. No. 2;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 9
US-09-248-796A-24357
; Sequence 24357, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

; SEQ ID NO 24357
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24357

Query Match 100.0%; Score 37; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 5.5;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 37 HHPHG 41

RESULT 10

US-09-902-540-13395
; Sequence 13395, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 13395
; LENGTH: 163
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-13395

Query Match 100.0%; Score 37; DB 4; Length 163;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 42 HHPHG 46

RESULT 11

US-09-270-767-61056
; Sequence 61056, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 61056
; LENGTH: 220
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-61056

Query Match 100.0%; Score 37; DB 4; Length 220;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 188 HHPHG 192

RESULT 12

US-09-540-236-3168
; Sequence 3168, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATARRHALIS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 3168
; LENGTH: 264
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-3168

Query Match 100.0%; Score 37; DB 4; Length 264;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 230 HHPHG 234

RESULT 13

US-09-949-016-6606
; Sequence 6606, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6606
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6606

Query Match 100.0%; Score 37; DB 4; Length 273;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 229 HHPHG 233

RESULT 14

US-09-949-016-9101
; Sequence 9101, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755

```
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9101
; LENGTH: 280
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9101

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 280;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 236 HHPHG 240

RESULT 15
US-09-252-991A-26545
; Sequence 26545, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26545
; LENGTH: 466
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26545

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 466;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 135 HHPHG 139

RESULT 16
US-09-976-594-64
; Sequence 64, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 64
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
```

```
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6673549 085596CD1
US-09-978-594-64

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 525;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 384 HHPHG 388

RESULT 17
US-09-919-039-62
; Sequence 62, Application US/09919039
; Patent No. 6727066
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR APPLICATION NUMBER: 60/222,113
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. 6727066 085596CD1
US-09-919-039-62

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 525;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 384 HHPHG 388

RESULT 18
US-09-518-959-8
; Sequence 8, Application US/09518959
; Patent No. 6548270
; GENERAL INFORMATION:
; APPLICANT: Dubin, Adrienne E
; APPLICANT: Erlander, Mark G
; APPLICANT: Huvar, Rene
; APPLICANT: Pyati, Jayashree
; TITLE OF INVENTION: DNA encoding human acid-sensing ion
; FILE REFERENCE: ORT-1197
; CURRENT APPLICATION NUMBER: US/09/518,959
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 8
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-959-8

Query Match
Best Local Similarity 100.0%; Score 37; DB 4; Length 539;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
```

```
Db          524 HHPHG 528

RESULT 19
US-09-518-959-9
; Sequence 9, Application US/09518959
; Patent No. 6548270
; GENERAL INFORMATION:
; APPLICANT: Dubin, Adrienne E
; APPLICANT: Erlander, Mark G
; APPLICANT: Huvar, Rene
; APPLICANT: Pyati, Jayashree
; TITLE OF INVENTION: DNA encoding human acid-sensing ion
; FILE REFERENCE: ORT-1197
; CURRENT APPLICATION NUMBER: US/09/518,959
; CURRENT FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-959-9

Query Match          100.0%; Score 37; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
   |||||
Db 524 HHPHG 528

RESULT 20
US-09-252-991A-23373
; Sequence 23373, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 23373
; LENGTH: 618
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-23373

Query Match          100.0%; Score 37; DB 4; Length 618;
Best Local Similarity 100.0%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
   |||||
Db 476 HHPHG 480

RESULT 21
US-09-248-796A-20281
; Sequence 20281, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
```

```
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 20281
; LENGTH: 810
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-20281

Query Match          100.0%; Score 37; DB 4; Length 810;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
   |||||
Db 50 HHPHG 54

RESULT 22
US-09-270-767-45546
; Sequence 45546, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45546
; LENGTH: 923
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-45546

Query Match          100.0%; Score 37; DB 4; Length 923;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
   |||||
Db 188 HHPHG 192

RESULT 23
US-09-546-013-29
; Sequence 29, Application US/09546013
; Patent No. 6610504
; GENERAL INFORMATION:
; APPLICANT: Yuan, Chong-Shen
; TITLE OF INVENTION: METHODS FOR ASSAYING S-ADENOSYLMETHIONINE-DEPENDENT Methyltransferase
; FILE REFERENCE: 10937-1652
; CURRENT APPLICATION NUMBER: US/09/546,013
; CURRENT FILING DATE: 2000-04-10
; EARLIER APPLICATION NUMBER: 09/347,878
; EARLIER FILING DATE: 1999-07-06
; EARLIER APPLICATION NUMBER: 09/457,205
; EARLIER FILING DATE: 1999-12-06
; NUMBER OF SEQ ID NOS: 94
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
```

OTHER INFORMATION: Description of Artificial Sequence: synthetic
OTHER INFORMATION: peptide representing metal-binding protein surface
OTHER INFORMATION: domains from human histidine rich glycoprotein
OTHER INFORMATION: (HRG)
PUBLICATION INFORMATION:
JOURNAL: J. Chromatogr.
VOLUME: 604
ISSUE: 1
PAGES: 125-132
DATE: 1992
US-09-546-013-29

Query Match 83.8%; Score 31; DB 4; Length 5;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 4
:||||
DB 2 HHPHG 5

RESULT 24
US-08-776-265-12
; Sequence 12, Application US/08776265
; Patent No. 6001631
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6001631el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4444
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; CLASIFICATION: 435

Query Match 83.8%; Score 31; DB 3; Length 7;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
:||||

Db 1 YHPHG 5
RESULT 25
US-09-398-184-12
; Sequence 12, Application US/09398184
; Patent No. 6649394
; GENERAL INFORMATION:
; APPLICANT: BLANCHE, Francis
; APPLICANT: CAMERON, Beatrice
; APPLICANT: CROUZET, Joel
; APPLICANT: FAMECHON, Alain
; APPLICANT: FERRERO, Lucia
; TITLE OF INVENTION: No. 6649394el Topoisomerase IV, Corresponding
; TITLE OF INVENTION: Nucleotide Sequences and Uses Thereof
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner, L.L.P.
; STREET: 1300 I. Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/398,184
; FILING DATE: 17-Sep-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/776,265
; FILING DATE: 24-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Einaudi, Carol P.
; REGISTRATION NUMBER: 32,220
; REFERENCE/DOCKET NUMBER: 03806.0394-00000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 408-4000
; TELEFAX: (202) 408-4444
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 12:
US-09-398-184-12

Query Match 83.8%; Score 31; DB 4; Length 7;
Best Local Similarity 80.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
:||||
DB 1 YHPHG 5

RESULT 26
US-08-159-339A-1005
; Sequence 1005, Application US/08159339A
; Patent No. 6037135
; GENERAL INFORMATION:
; APPLICANT: Kubo, Ralph T.
; APPLICANT: Grey, Howard M.
; APPLICANT: Sette, Alessandro
; APPLICANT: Celis, Eteban
; TITLE OF INVENTION: HLA Binding peptides and Their
; TITLE OF INVENTION: Uses

NUMBER OF SEQUENCES: 1254
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: CA
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/159,339A
FILING DATE: 29-NOV-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/926,666
FILING DATE: 07-AUG-1992
APPLICATION NUMBER: US 08/027,746
FILING DATE: 05-MAR-1993
APPLICATION NUMBER: US 08/103,396
FILING DATE: 06-AUG-1993
ATTORNEY/AGENT INFORMATION:
NAME: Weber, Ellen Lauver
REGISTRATION NUMBER: 32,762
REFERENCE/DOCKET NUMBER: 018623-005030US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
TELEX:
INFORMATION FOR SEQ ID NO: 1005:
SEQUENCE CHARACTERISTICS:
LENGTH: 9 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-159-339A-1005

Query Match 83.8%; Score 31; DB 3; Length 9;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
DB 3 HHPH 6

RESULT 27
US-09-082-092-14
; Sequence 14, Application US/09082092
; Patent No. 6251628
; GENERAL INFORMATION:
; APPLICANT: Nakao, Atsuhito
; APPLICANT: Moren, Anita
; APPLICANT: Heuchel, Rainer
; APPLICANT: Itoh, Susumu
; APPLICANT: Afrakhte, Mozghan
; APPLICANT: Soucheinytskyi, Serhiy
; APPLICANT: Brodin, Greger
; APPLICANT: Landstrom, Marene
; APPLICANT: Heldin, Nils-Erik
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten Dijke, Peter
; TITLE OF INVENTION: SMAD7 AND USES THEREOF
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
; STREET: 600 Atlantic Avenue
; CITY: Boston
; STATE: MA

COUNTRY: U.S.A.
ZIP: 02210-2211
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/082,092
FILING DATE: 20-MAY-1998
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/047,221
FILING DATE: 20-MAY-1997
APPLICATION NUMBER: 60/060,465
FILING DATE: 30-SEP-1997
APPLICATION NUMBER: 60/075,940
FILING DATE: 25-FEB-1998
APPLICATION NUMBER: 60/077,033
FILING DATE: 06-MAR-1998
ATTORNEY/AGENT INFORMATION:
NAME: Van Amsterdam, John R.
REGISTRATION NUMBER: 40,212
REFERENCE/DOCKET NUMBER: L0461/7032
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-720-3500
TELEFAX: 617-720-2441
TELEX:
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 15 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-082-092-14

Query Match 83.8%; Score 31; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
DB 10 HHPH 13

RESULT 28
US-09-885-722A-14
; Sequence 14, Application US/09885722A
; Patent No. 6605443
; GENERAL INFORMATION:
; APPLICANT: Nakao, Atsuhito
; APPLICANT: Heldin, Carl-Henrik
; APPLICANT: ten Dijke, Peter
; TITLE OF INVENTION: SMAD7 AND USES THEREOF
; FILE REFERENCE: L00461.70117.US
; CURRENT APPLICATION NUMBER: US/09/885,722A
; CURRENT FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: US 09/082,092
; PRIOR FILING DATE: 1998-05-20
; PRIOR APPLICATION NUMBER: US 60/077,033
; PRIOR FILING DATE: 1998-03-06
; PRIOR APPLICATION NUMBER: US 60/075,940
; PRIOR FILING DATE: 1998-02-25
; PRIOR APPLICATION NUMBER: US 60/060,465
; PRIOR FILING DATE: 1997-09-30
; PRIOR APPLICATION NUMBER: US 60/047,221
; PRIOR FILING DATE: 1997-05-20
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 14
; LENGTH: 15
; TYPE: PRT

```

; ORGANISM: Mus musculus
US-09-885-722A-14

Query Match      83.8%; Score 31; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      10 HHPH 13

RESULT 29
US-09-227-357-575
; Sequence 575, Application US/09227357
; Patent No. 6342581
; GENERAL INFORMATION:
; APPLICANT: Fischer et al.
; TITLE OF INVENTION: 123 Human Secreted Proteins
; FILE REFERENCE: P2010P1
; CURRENT APPLICATION NUMBER: US/09/227,357
; CURRENT FILING DATE: 1999-01-08
; EARLIER APPLICATION NUMBER: PCT/US98/13684
; EARLIER FILING DATE: 1998-07-07
; EARLIER APPLICATION NUMBER: 60/051,926
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,793
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,925
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,929
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,803
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,732
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,931
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,932
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,916
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,930
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,918
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,920
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,733
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/052,795
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,919
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/051,928
; EARLIER FILING DATE: 1997-07-08
; EARLIER APPLICATION NUMBER: 60/055,722
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,723
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,948
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,949
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,953
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,950
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,947
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/055,964
; EARLIER FILING DATE: 1997-08-18
; EARLIER APPLICATION NUMBER: 60/056,360

; ORGANISM: Mus musculus
US-09-885-722A-14

Query Match      83.8%; Score 31; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      10 HHPH 13

RESULT 30
US-09-910-009A-158
; Sequence 158, Application US/09910009A
; Patent No. 6727226
; GENERAL INFORMATION:
; APPLICANT: University of Utah Research Foundation
; APPLICANT: Cognetix, Inc.
; APPLICANT: Olivera, Baldomero M.
; APPLICANT: McIntosh, J. Michael
; APPLICANT: Garrett, James E.
; APPLICANT: Watkins, Maren
; APPLICANT: Cruz, Lourdes J.
; APPLICANT: Shon, Ki-Joon
; APPLICANT: Jacobsen, Richard
; APPLICANT: Jones, Robert M.
; APPLICANT: Cartier, G. Edward
; APPLICANT: Shen, Greg S.
; APPLICANT: Wagstaff, John D.
; TITLE OF INVENTION: Mu-Conopeptides
; FILE REFERENCE: 2314-242
; CURRENT APPLICATION NUMBER: US/09/910,009A
; CURRENT FILING DATE: 2001-07-23
; PRIOR APPLICATION NUMBER: US 60/219,619
; PRIOR FILING DATE: 2000-07-21
; PRIOR APPLICATION NUMBER: US 60/245,157
; PRIOR FILING DATE: 2000-11-03
; PRIOR APPLICATION NUMBER: US 60/264,319
; PRIOR FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: US 60/277,270
; PRIOR FILING DATE: 2001-03-21
; NUMBER OF SEQ ID NOS: 520
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 158
; LENGTH: 69
; TYPE: PRT
; ORGANISM: Conus marmoreus
US-09-910-009A-158

Query Match      83.8%; Score 31; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY      1 HHPH 4
Db      44 HHPH 47

RESULT 31
US-09-716-129-73
; Sequence 73, Application US/09716129
; Patent No. 6632920
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 36 Human Secreted Proteins
; FILE REFERENCE: P2025P1
; CURRENT APPLICATION NUMBER: US/09/716,129
; CURRENT FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/076,053
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,057
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,052
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,054
; PRIOR FILING DATE: 1998-02-26
; PRIOR APPLICATION NUMBER: 60/076,051
; PRIOR FILING DATE: 1998-02-26
; NUMBER OF SEQ ID NOS: 186
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 73
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (75)
; OTHER INFORMATION: Xaa equals stop translation
US-09-716-129-73

Query Match      83.8%; Score 31; DB 4; Length 75;
Best Local Similarity 100.0%; Pred.No. 70;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      23 HHPH 26

RESULT 32
US-09-489-039A-10688
; Sequence 10688, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10688
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-10688

Query Match      83.8%; Score 31; DB 4; Length 76;
Best Local Similarity 100.0%; Pred.No. 71;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      44 HHPH 47
```

```
Db      72 HHPH 75

RESULT 33
US-09-270-767-57094
; Sequence 57094, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57094
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57094

Query Match      83.8%; Score 31; DB 4; Length 84;
Best Local Similarity 100.0%; Pred.No. 79;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      8 HHPH 11

RESULT 34
US-09-270-767-36032
; Sequence 36032, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 36032
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-36032

Query Match      83.8%; Score 31; DB 4; Length 91;
Best Local Similarity 100.0%; Pred.No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPH 4
Db      3 HHPH 6

RESULT 35
US-09-270-767-51249
; Sequence 51249, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 51249
; LENGTH: 91
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
```

US-09-270-767-51249

Query Match 83.8%; Score 31; DB 4; Length 91;
Best Local Similarity 100.0%; Pred. No. 85;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
DB 3 HHPH 6

RESULT 36

US-09-248-796A-23534
; Sequence 23534, Application US/09248796A
; Patent No. 6747137

; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.132

; CURRENT APPLICATION NUMBER: US/09/248,796A

; CURRENT FILING DATE: 1999-02-12

; PRIOR APPLICATION NUMBER: US 60/074,725

; PRIOR FILING DATE: 1998-02-13

; PRIOR APPLICATION NUMBER: US 60/096,409

; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208

; SEQ ID NO 23534

; LENGTH: 92

; TYPE: PRT

; ORGANISM: Candida albicans

US-09-248-796A-23534

; Query Match

83.8%; Score 31; DB 4; Length 92;
Best Local Similarity 80.0%; Pred. No. 86;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||:|
DB 73 HHPYG 77

RESULT 37

US-09-640-211A-653

; Sequence 653, Application US/09640211A

; Patent No. 6833446

; GENERAL INFORMATION:

; APPLICANT: Wood, Marion

; APPLICANT: Shenk, Michael A.

; APPLICANT: McGrath, Annette

; APPLICANT: Glenn, Matthew

; TITLE OF INVENTION: Compositions and Methods for the

; TITLE OF INVENTION: Modification of Gene Transcription

; FILE REFERENCE: 11000.1021C1U

; CURRENT APPLICATION NUMBER: US/09/640,211A

; CURRENT FILING DATE: 2000-08-16

; NUMBER OF SEQ ID NOS: 2368

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 653

; LENGTH: 99

; TYPE: PRT

; ORGANISM: Eucalyptus grandis

US-09-640-211A-653

Query Match

83.8%; Score 31; DB 4; Length 99;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
DB 9 HHPH 12

RESULT 38

US-09-270-767-58449

; Sequence 58449, Application US/09270767

; Patent No. 6703491

; GENERAL INFORMATION:

; APPLICANT: Homburger et al.

; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster

; FILE REFERENCE: File Reference: 7326-094

; CURRENT APPLICATION NUMBER: US/09/270,767

; CURRENT FILING DATE: 1999-03-17

; NUMBER OF SEQ ID NOS: 62517

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 58449

; LENGTH: 102

; TYPE: PRT

; ORGANISM: Drosophila melanogaster

; FEATURE:

; OTHER INFORMATION: Xaa means any amino acid

US-09-270-767-58449

Query Match

83.8%; Score 31; DB 4; Length 102;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
DB 87 HHPH 90

RESULT 39

US-09-461-325-160

; Sequence 160, Application US/09461325A

; Patent No. 6475753

; GENERAL INFORMATION:

; APPLICANT: Ruben et al.

; TITLE OF INVENTION: 94 Human Secreted Proteins

; FILE REFERENCE: P2029Pl

; CURRENT APPLICATION NUMBER: US/09/461,325A

; CURRENT FILING DATE: 1999-12-14

; EARLIER APPLICATION NUMBER: PCT/US99/13418

; EARLIER FILING DATE: 1999-06-15

; EARLIER APPLICATION NUMBER: 60/089,507

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,508

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,509

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/089,510

; EARLIER FILING DATE: 1998-06-16

; EARLIER APPLICATION NUMBER: 60/090,112

; EARLIER FILING DATE: 1998-06-22

; EARLIER APPLICATION NUMBER: 60/090,113

; EARLIER FILING DATE: 1998-06-22

; NUMBER OF SEQ ID NOS: 532

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 160

; LENGTH: 115

; TYPE: PRT

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (96)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (100)

; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

US-09-461-325-160

Query Match

83.8%; Score 31; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|
Db 104 HHPH 107

RESULT 40
US-10-012-542-160
; Sequence 160, Application US/10012542
; Patent No. 6627741
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 94 Human Secreted Proteins
; FILE REFERENCE: P2029P1
; CURRENT APPLICATION NUMBER: US/10/012,542
; CURRENT FILING DATE: 2001-12-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/461,325
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-12-14
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,507
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,508
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,509
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/089,510
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-16
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,112
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 60/090,113
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-06-22
; NUMBER OF SEQ ID NOS: 532
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 160
; LENGTH: 115
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (96)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (100)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-012-542-160

Query Match 83.8%; Score 31; DB 4; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|
Db 104 HHPH 107

Search completed: June 15, 2005, 14:24:19
Job time : 24.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-8
Perfect score: 37
Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	5	14	US-10-074-225A-8
2	37	100.0	6	14	US-10-153-312A-1
3	37	100.0	11	9	US-09-809-657-1
4	37	100.0	11	9	US-09-848-512A-1
5	37	100.0	11	9	US-09-742-494-1
6	37	100.0	11	14	US-10-153-312A-2
7	37	100.0	16	14	US-10-153-312A-3
8	37	100.0	26	9	US-09-809-657-2
9	37	100.0	26	9	US-09-848-512A-2
10	37	100.0	26	9	US-09-742-494-2
11	37	100.0	26	14	US-10-153-312A-4

12	37	100.0	84	15	US-10-424-599-277019	Sequence 277019,
13	37	100.0	97	14	US-10-106-698-4437	Sequence 4437, Ap
14	37	100.0	101	14	US-10-074-225A-6	Sequence 6, Appli
15	37	100.0	116	16	US-10-437-963-149790	Sequence 149790,
16	37	100.0	129	16	US-10-425-115-324288	Sequence 324288,
17	37	100.0	138	16	US-10-437-963-142824	Sequence 142824,
18	37	100.0	148	14	US-10-074-225A-5	Sequence 5, Appli
19	37	100.0	151	9	US-09-864-761-42051	Sequence 42051, A
20	37	100.0	152	16	US-10-437-963-128472	Sequence 128472,
21	37	100.0	185	15	US-10-425-114-58759	Sequence 58759, A
22	37	100.0	259	15	US-10-282-1236-62936	Sequence 62936, A
23	37	100.0	273	10	US-09-981-151A-64	Sequence 64, Appl
24	37	100.0	273	10	US-09-981-151A-65	Sequence 65, Appl
25	37	100.0	273	15	US-10-295-027-344	Sequence 344, App
26	37	100.0	273	15	US-10-295-027-1250	Sequence 1250, Ap
27	37	100.0	273	15	US-10-072-012-593	Sequence 593, App
28	37	100.0	273	15	US-10-072-012-596	Sequence 185, App
29	37	100.0	273	15	US-10-188-832-185	Sequence 185, App
30	37	100.0	350	15	US-10-369-493-19871	Sequence 19871, A
31	37	100.0	460	15	US-10-424-599-205101	Sequence 205101,
32	37	100.0	492	16	US-10-808-807-8	Sequence 8, Appli
33	37	100.0	525	10	US-09-919-039-62	Sequence 62, Appl
34	37	100.0	525	14	US-10-074-225A-1	Sequence 1, Appli
35	37	100.0	525	17	US-10-868-577A-42	Sequence 42, Appl
36	37	100.0	526	14	US-10-074-225A-3	Sequence 3, Appli
37	37	100.0	539	10	US-09-772-180A-2	Sequence 2, Appli
38	37	100.0	539	10	US-09-772-180A-8	Sequence 8, Appli
39	37	100.0	539	15	US-10-295-027-290	Sequence 290, App
40	37	100.0	587	10	US-09-772-180A-4	Sequence 4, Appli
41	37	100.0	19695	15	US-10-084-846A-3	Sequence 3, Appli
42	31	83.8	5	14	US-10-419-802-1	Sequence 1, Appli
43	31	83.8	5	15	US-10-373-238-25	Sequence 25, Appl
44	31	83.8	7	15	US-10-646-807-12	Sequence 12, Appl
45	31	83.8	10	17	US-10-883-760-24	Sequence 24, Appl

ALIGNMENTS

RESULT 1

US-10-074-225A-8
; Sequence 8, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; PRIOR FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-8

Query Match 100.0%; Score 37; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 1 HHPHG 5

```

RESULT 2
US-10-153-312A-1
; Sequence 1, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086 US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 6
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-1

Query Match      100.0%; Score 37; DB 14; Length 6;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 3
US-09-809-657-1
; Sequence 1, Application US/09809657
; Patent No. US20010014479A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USF/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 4
US-09-848-512A-1
; Sequence 1, Application US/09848512A
; Patent No. US20020037517A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS FOR SEQUENCING BIOPOLYMERS

```

```

; FILE REFERENCE: P00798USG/09306611
; CURRENT APPLICATION NUMBER: US/09/848,512A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-848-512A-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 5
US-09-742-494-1
; Sequence 1, Application US/09742494
; Patent No. US20020155620A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-1

Query Match      100.0%; Score 37; DB 9; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      2 HHPHG 6

RESULT 6
US-10-153-312A-2
; Sequence 2, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 11
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide

```

US-10-153-312A-2

Query Match 100.0%; Score 37; DB 14; Length 11;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 2 HHPHG 6

RESULT 7

US-10-153-312A-3
; Sequence 3, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24
; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 16
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide

US-10-153-312A-3

Query Match 100.0%; Score 37; DB 14; Length 16;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 2 HHPHG 6

RESULT 8

US-09-809-657-2
; Sequence 2, Application US/09809657
; Patent No. US20010014479A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/809,657
; CURRENT FILING DATE: 2001-03-15
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-809-657-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 2 HHPHG 6

RESULT 9

US-09-848-512A-2
; Sequence 2, Application US/09848512A
; Patent No. US20020037517A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS FOR SEQUENCING BIOPOLYMERS
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/848,512A
; CURRENT FILING DATE: 2000-05-03
; PRIOR APPLICATION NUMBER: 08/068,896
; PRIOR FILING DATE: 1993-05-28
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-848-512A-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 2 HHPHG 6

RESULT 10

US-09-742-494-2
; Sequence 2, Application US/09742494
; Patent No. US20020155620A1
; GENERAL INFORMATION:
; APPLICANT: HUTCHENS, WILLIAM
; APPLICANT: YIP, TAI-TUNG
; TITLE OF INVENTION: METHODS AND APPARATUS FOR DESORPTION AND IONIZATION OF ANALYTES
; FILE REFERENCE: P00798USE/09306611
; CURRENT APPLICATION NUMBER: US/09/742,494
; CURRENT FILING DATE: 2000-12-20
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-742-494-2

Query Match 100.0%; Score 37; DB 9; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 2 HHPHG 6

RESULT 11

US-10-153-312A-4
; Sequence 4, Application US/10153312A
; Publication No. US20030040016A1
; GENERAL INFORMATION:
; APPLICANT: Singh, Sharat
; APPLICANT: Zivin, Robert Allan
; TITLE OF INVENTION: Analyzing Phosphorylated Proteins
; FILE REFERENCE: 50225-8086.US03
; CURRENT APPLICATION NUMBER: US/10/153,312A
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/334,902
; PRIOR FILING DATE: 2001-10-24

; PRIOR APPLICATION NUMBER: US 60/292,548
; PRIOR FILING DATE: 2001-05-21
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: metal binding peptide
US-10-153-312A-4

Query Match 100.0%; Score 37; DB 14; Length 26;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 2 HHPHG 6

RESULT 12

US-10-424-599-277019
; Sequence 277019, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 277019
; LENGTH: 84
; TYPE: PRT

; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_92167C.1.pap
US-10-424-599-277019

Query Match 100.0%; Score 37; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 52 HHPHG 56

RESULT 13

US-10-106-698-4437
; Sequence 4437, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.

; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03

; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: PatentIn Ver. 3.0
; SEQ ID NO 4437
; LENGTH: 97
; TYPE: PRT

; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: MISC_FEATURE
; LOCATION: (26)_____
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (29)_____
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC_FEATURE
; LOCATION: (33)_____
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4437

Query Match 100.0%; Score 37; DB 14; Length 97;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 53 HHPHG 57

RESULT 14

US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match 100.0%; Score 37; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 13 HHPHG 17

RESULT 15

US-10-437-963-149790
; Sequence 149790, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping

; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 149790
; LENGTH: 116
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_50089C.1.pap
US-10-437-963-149790

Query Match 100.0%; Score 37; DB 16; Length 116;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 89 HHPHG 93

RESULT 16

US-10-425-115-324288
; Sequence 324288, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 324288
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58924C.1.pap
US-10-425-115-324288

Query Match 100.0%; Score 37; DB 16; Length 129;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 85 HHPHG 89

RESULT 17

US-10-437-963-142824
; Sequence 142824, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14

; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 142824
; LENGTH: 138
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_43793C.1.pap
US-10-437-963-142824

Query Match 100.0%; Score 37; DB 16; Length 138;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 36 HHPHG 40

RESULT 18

US-10-074-225A-5
; Sequence 5, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 148
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-5

Query Match 100.0%; Score 37; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 35 HHPHG 39

RESULT 19

US-09-864-761-42051
; Sequence 42051, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Weneheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04

; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 42051
; LENGTH: 151
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AL122010.2
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.4
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.7
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.2
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 3.4
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EST HUMAN HIT: BE182360.1, EVALUE 3.00e-64
; OTHER INFORMATION: SWISSPROT HIT: O95377, EVALUE 2.00e-89
US-09-864-761-42051

Query Match 100.0%; Score 37; DB 9; Length 151;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 128 HHPHG 132

RESULT 20
US-10-437-963-128472
; Sequence 128472, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B

; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128472
; LENGTH: 152
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30821C.1.pep
US-10-437-963-128472

Query Match 100.0%; Score 37; DB 16; Length 152;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 24 HHPHG 28

RESULT 21
US-10-425-114-58759
; Sequence 58759, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 58759
; LENGTH: 185
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700214339_FLI.pep
US-10-425-114-58759

Query Match 100.0%; Score 37; DB 15; Length 185;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
Db 117 HHPHG 121

RESULT 22
US-10-282-122A-62936
; Sequence 62936, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haeelbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A

```

; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 62936
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
US-10-282-122A-62936

; Query Match 100.0%; Score 37; DB 15; Length 259;
; Best Local Similarity 100.0%; Pred. No. 2.6e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 229 HHPHG 233

RESULT 23
US-09-981-151A-64
; Sequence 64, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zernusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17

```

```

; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,243
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/242,152
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/242,482
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,611
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,612
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/242,880
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: 60/242,881
; PRIOR FILING DATE: 2000-10-24
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 160
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-981-151A-64

; Query Match 100.0%; Score 37; DB 10; Length 273;
; Best Local Similarity 100.0%; Pred. No. 2.7e+02;
; Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 229 HHPHG 233

RESULT 24
US-09-981-151A-65
; Sequence 65, Application US/09981151A
; Publication No. US20030212256A1
; GENERAL INFORMATION:
; APPLICANT: Edinger, Shlomit R
; APPLICANT: Gerlach, Valerie
; APPLICANT: MacDougall, John R
; APPLICANT: Malyankar, Muriel M
; APPLICANT: Smithson, Glenda
; APPLICANT: Millet, Isabelle
; APPLICANT: Peyman, John A
; APPLICANT: Stone, David J
; APPLICANT: Gunther, Erik
; APPLICANT: Ellerman, Karen
; APPLICANT: Shinkets, Richard A
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Guo, Xiaojia
; APPLICANT: Patturajan, Meera
; APPLICANT: Taupier Jr, Raymond J
; APPLICANT: Burgess, Catherine E
; APPLICANT: Zernusen, Bryan D
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Spytek, Kimberly A
; APPLICANT: Gangolli, Esha A
; APPLICANT: Fernandes, Elma R
; APPLICANT: Gorman, Linda
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-168
; CURRENT APPLICATION NUMBER: US/09/981,151A
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 60/241,040
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,058
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/241,063
; PRIOR FILING DATE: 2000-10-17

```

```
/ PRIOR APPLICATION NUMBER: 60/241,243
/ PRIOR FILING DATE: 2000-10-17
/ PRIOR APPLICATION NUMBER: 60/242,152
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/242,482
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,611
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,612
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/242,880
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: 60/242,881
/ PRIOR FILING DATE: 2000-10-24
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 160
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 65
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: Homo sapiens
/ NAME/KEY: VARIANT
/ LOCATION: (41)
/ OTHER INFORMATION: Wherein Xaa is any amino acid.
US-09-981-151A-65

Query Match      100.0%; Score 37; DB 10; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      229 HHPHG 233

RESULT 25
US-10-295-027-344
/ Sequence 344, Application US/10295027
/ Publication No. US2003023350A1
/ GENERAL INFORMATION:
/ APPLICANT: Afar, Daniel
/ APPLICANT: Aziz, Natasha
/ APPLICANT: Ginsberg, Wendy M.
/ APPLICANT: Gish, Kurt C.
/ APPLICANT: Glynn, Richard
/ APPLICANT: Hevezi, Peter A.
/ APPLICANT: Mack, David H.
/ APPLICANT: Murray, Richard
/ APPLICANT: Watson, Susan R.
/ APPLICANT: Eos Biotechnology, Inc.
/ TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
/ FILE REFERENCE: 018501-012500US
/ CURRENT APPLICATION NUMBER: US/10/295,027
/ CURRENT FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: US 09/663,733
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: US 60/350,666
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: US 60/335,394
/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1250
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-295-027-1250

Query Match      100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      229 HHPHG 233

RESULT 26
US-10-295-027-1250
/ Sequence 1250, Application US/10295027
/ Publication No. US2003023350A1
/ GENERAL INFORMATION:
/ APPLICANT: Afar, Daniel
/ APPLICANT: Aziz, Natasha
/ APPLICANT: Ginsberg, Wendy M.
/ APPLICANT: Gish, Kurt C.
/ APPLICANT: Glynn, Richard
/ APPLICANT: Hevezi, Peter A.
/ APPLICANT: Mack, David H.
/ APPLICANT: Murray, Richard
/ APPLICANT: Watson, Susan R.
/ APPLICANT: Eos Biotechnology, Inc.
/ TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
/ FILE REFERENCE: 018501-012500US
/ CURRENT APPLICATION NUMBER: US/10/295,027
/ CURRENT FILING DATE: 2002-11-13
/ PRIOR APPLICATION NUMBER: US 09/663,733
/ PRIOR FILING DATE: 2000-09-15
/ PRIOR APPLICATION NUMBER: US 60/350,666
/ PRIOR FILING DATE: 2001-11-13
/ PRIOR APPLICATION NUMBER: US 60/335,394
/ PRIOR FILING DATE: 2001-11-15
/ PRIOR APPLICATION NUMBER: US 60/332,464
/ PRIOR FILING DATE: 2001-11-21
/ PRIOR APPLICATION NUMBER: US 60/334,393
/ PRIOR FILING DATE: 2001-11-29
/ PRIOR APPLICATION NUMBER: US 60/340,376
/ PRIOR FILING DATE: 2001-12-14
/ PRIOR APPLICATION NUMBER: US 60/347,211
/ PRIOR FILING DATE: 2002-01-08
/ PRIOR APPLICATION NUMBER: US 60/347,349
/ PRIOR FILING DATE: 2002-01-10
/ PRIOR APPLICATION NUMBER: US 60/355,250
/ PRIOR FILING DATE: 2002-02-08
/ PRIOR APPLICATION NUMBER: US 60/356,714
/ PRIOR FILING DATE: 2002-02-13
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 1386
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 1250
/ LENGTH: 273
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-295-027-1250

Query Match      100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 1 HHPHG 5
|||||
Db 229 HHPHG 233

RESULT 27
US-10-072-012-593
; Sequence 593, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/266,975
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 593
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (41)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-072-012-593

Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 229 HHPHG 233

RESULT 28
US-10-072-012-596
; Sequence 596, Application US/10072012
; Publication No. US20040033493A1
; GENERAL INFORMATION:
; APPLICANT: Tchernev, Velizar
; APPLICANT: Spytek, Kimberly
; APPLICANT: Zerhusen, Bryan
; APPLICANT: Patturajan, Meera
; APPLICANT: Shimkete, Richard
; APPLICANT: Li, Li
; APPLICANT: Gangolli, Esha
; APPLICANT: Padigar, Muralidhara
; APPLICANT: Anderson, David W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Miller, Charles E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Taupier Jr, Raymond J.
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Colman, Steven D.
; APPLICANT: Wolenc, Adam R.
; APPLICANT: Pena, Carol E. A
; APPLICANT: Furtak, Katarzyna
; APPLICANT: Grosse, William M.
; APPLICANT: Alsobrook II, John P.
; APPLICANT: Lepley, Denise M.
; APPLICANT: Rieger, Daniel K.
; APPLICANT: Burgess, Catherine E.
; TITLE OF INVENTION: Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-258
; CURRENT APPLICATION NUMBER: US/10/072,012
; PRIOR FILING DATE: 2002-01-31
; PRIOR APPLICATION NUMBER: 60/265,102
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: 60/265,514
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,517
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,412
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/265,395
; PRIOR FILING DATE: 2001-01-31
; PRIOR APPLICATION NUMBER: 60/266,406
; PRIOR FILING DATE: 2001-02-02
; PRIOR APPLICATION NUMBER: 60/266,767
; PRIOR FILING DATE: 2001-02-05
; PRIOR APPLICATION NUMBER: 60/267,057
; PRIOR FILING DATE: 2001-02-07
; PRIOR APPLICATION NUMBER: 60/267,459
; PRIOR FILING DATE: 2001-02-08
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1391
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 596
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: VARIANT
; LOCATION: (41)
; OTHER INFORMATION: Wherein Xaa is any amino acid.
US-10-072-012-596

Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

Db 229 HHPHG 233
|||||

RESULT 29
US-10-188-832-185
; Sequence 185, Application US/10188832
; Publication No. US20040076955A1
; GENERAL INFORMATION:
; APPLICANT: Mack, David H.
; APPLICANT: Aziz, Natasha
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Bladder Cancer, Compositions
; TITLE OF INVENTION: and Methods of Screening for Modulators of Bladder
; TITLE OF INVENTION: Cancer
; FILE REFERENCE: 018501-002330US
; CURRENT APPLICATION NUMBER: US/10/188,832
; CURRENT FILING DATE: 2002-11-22
; PRIOR APPLICATION NUMBER: US 60/302,814
; PRIOR FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: US 60/310,099
; PRIOR FILING DATE: 2001-08-03
; PRIOR APPLICATION NUMBER: US 60/343,705
; PRIOR FILING DATE: 2001-11-08
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/372,246
; PRIOR FILING DATE: 2002-04-12
; NUMBER OF SEQ ID NOS: 207
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 185
; LENGTH: 273
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-832-185

Query Match 100.0%; Score 37; DB 15; Length 273;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 229 HHPHG 233

RESULT 30
US-10-369-493-19871
; Sequence 19871, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 19871
; LENGTH: 350
; TYPE: PRT
; ORGANISM: No. US20030233675A1toc punctiforme
US-10-369-493-19871

Query Match 100.0%; Score 37; DB 15; Length 350;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 174 HHPHG 178

RESULT 31
US-10-424-599-205101
; Sequence 205101, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 205101
; LENGTH: 460
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(460)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_27235C.1.pap
US-10-424-599-205101

Query Match 100.0%; Score 37; DB 15; Length 460;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 15 HHPHG 19

RESULT 32
US-10-808-807-8
; Sequence 8, Application US/10808807
; Publication No. US20040253663A1
; GENERAL INFORMATION:
; APPLICANT: E.I. du Pont deNemours and Co., Inc.
; APPLICANT: Cheng, Qiong
; APPLICANT: Tao, Luan
; APPLICANT: Sedkova, Natalia
; TITLE OF INVENTION: GENES ENCODING CAROTENOID COMPOUNDS
; FILE REFERENCE: CL2365 US NA
; CURRENT APPLICATION NUMBER: US/10/808,807
; CURRENT FILING DATE: 2004-03-24
; PRIOR APPLICATION NUMBER: US 60/477,874
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 60/527,083
; PRIOR FILING DATE: 2003-12-03
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Pantoea agglomerans strain DC404
US-10-808-807-8

Query Match 100.0%; Score 37; DB 16; Length 492;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 286 HHPHG 290

```
RESULT 33
US-09-919-039-62
; Sequence 62, Application US/09919039
; Publication No. US20030108871A1
; GENERAL INFORMATION:
; APPLICANT: Kaser, Matthew R.
; TITLE OF INVENTION: GENES EXPRESSED IN TREATED HUMAN C3A LIVER CELL CULTURES
; FILE REFERENCE: PA-0035 US
; CURRENT APPLICATION NUMBER: US/09/919,039
; CURRENT FILING DATE: 2002-09-09
; PRIOR FILING DATE: 2002-09-09
; PRIOR FILING DATE: 2002-11-13
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 401
; SOFTWARE: PERL Program
; SEQ ID NO 62
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030108871A1 085596CD1
US-09-919-039-62

Query Match 100.0%; Score 37; DB 10; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 384 HHPHG 388

RESULT 34
US-10-074-225A-1
; Sequence 1, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR FILING DATE: 2001-02-14
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-074-225A-1

Query Match 100.0%; Score 37; DB 14; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 384 HHPHG 388

RESULT 35
US-10-868-577A-42
; Sequence 42, Application US/10868577A
; Publication No. US20050032697A1
; GENERAL INFORMATION:
; APPLICANT: Alitalo et al.

; TITLE OF INVENTION: HEPARIN BINDING VEGFR-3 LIGANDS
; FILE REFERENCE: 28967/39359A
; CURRENT APPLICATION NUMBER: US/10/868,577A
; CURRENT FILING DATE: 2004-06-14
; PRIOR APPLICATION NUMBER: US 60/478,390
; PRIOR FILING DATE: 2003-06-12
; PRIOR APPLICATION NUMBER: US 10/669,176
; PRIOR FILING DATE: 2003-09-23
; NUMBER OF SEQ ID NOS: 69
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 525
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-868-577A-42

Query Match 100.0%; Score 37; DB 17; Length 525;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 384 HHPHG 388

RESULT 36
US-10-074-225A-3
; Sequence 3, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match 100.0%; Score 37; DB 14; Length 526;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 333 HHPHG 337

RESULT 37
US-09-772-180A-2
; Sequence 2, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848
```

; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 539
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-772-180A-2

Query Match 100.0%; Score 37; DB 10; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
Db 524 HHPHG 528

RESULT 38
US-09-772-180A-8
; Sequence 8, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 8
; LENGTH: 539
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
US-09-772-180A-8

Query Match 100.0%; Score 37; DB 10; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
Db 524 HHPHG 528

RESULT 39
US-10-295-027-290
; Sequence 290, Application US/10295027
; Publication No. US2003032350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard

; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE OF INVENTION: Methods of Screening for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 290
; LENGTH: 539
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-290

Query Match 100.0%; Score 37; DB 15; Length 539;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 HHPHG 5
Db 524 HHPHG 528

RESULT 40
US-09-772-180A-4
; Sequence 4, Application US/09772180A
; Publication No. US20030027749A1
; GENERAL INFORMATION:
; APPLICANT: David C. Harrison
; APPLICANT: John Davis
; APPLICANT: Sharon Bingham
; APPLICANT: Trudy R. Doe
; APPLICANT: Simon Topp
; TITLE OF INVENTION: NOVEL COMPOUNDS
; FILE REFERENCE: GH-30021-C1
; CURRENT APPLICATION NUMBER: US/09/772,180A
; CURRENT FILING DATE: 2001-01-29
; PRIOR APPLICATION NUMBER: 09/063,848
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 9708936.1
; PRIOR FILING DATE: 1997-05-01
; PRIOR APPLICATION NUMBER: 97310289.0
; PRIOR FILING DATE: 1997-12-18
; PRIOR APPLICATION NUMBER: 9803566.0
; PRIOR FILING DATE: 1998-02-19
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 4

```

; LENGTH: 587
; TYPE: PRT
; ORGANISM: HOMO SAPIENS
; FEATURE: FEATURE
; NAME/KEY: UNSURE
; LOCATION: (328) (356) (357) (358) (359) (360) (361) (362) (363) (364) (365) (366)
; LOCATION: (367) (368) (369) (370) (371) (372) (373) (374) (375) (376) (377) (378)
; LOCATION: (379) (380) (381) (382) (383) (384) (385) (386) (387) (388) (389) (390)
; LOCATION: (391) (392) (393) (394) (395) (396) (397) (398) (399) (400) (401) (402)
; LOCATION: (403) (404) (405) (406) (407) (408) (409) (410) (411) (412) (413) (414)
; LOCATION: (415) (416) (417)
; OTHER INFORMATION: OTHER INFORMATION: Partial Amino Acid Sequence
US-09-772-180A-4

Query Match      100.0%; Score 37; DB 10; Length 587;
Best Local Similarity 100.0%; Pred. No. 5.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      572 HHPHG 576

Search completed: June 15, 2005, 14:50:18
Job time : 80.75 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-8

Perfect score: 37

Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl2:*

3: Pirl3:*

4: Pirl4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	100.0	419	1 S70617	homeotic protein g
2	37	100.0	445	2 A60488	histidine-rich gly
3	37	100.0	491	2 H83979	lysine decarboxyla
4	37	100.0	525	1 KGHUGH	histidine-rich gly
5	37	100.0	602	2 D75618	hypothetical prote
6	37	100.0	721	2 E70766	hypothetical prote
7	37	100.0	753	1 D72660	probable aldehyde
8	37	100.0	1441	2 T13717	CRAG protein - fru
9	31	83.8	76	2 F97793	hypothetical prote
10	31	83.8	84	2 AH3162	IS21 family transp
11	31	83.8	94	2 A49832	DNA gyrase A - Sta
12	31	83.8	95	2 S49447	DNA topoisomerase
13	31	83.8	95	2 S54254	DNA topoisomerase
14	31	83.8	105	2 S62940	probable membrane
15	31	83.8	110	2 F69760	hypothetical prote
16	31	83.8	113	2 E82687	cytochrome O ubiq
17	31	83.8	130	2 S14983	extensin class I (
18	31	83.8	134	2 D72634	hypothetical prote
19	31	83.8	136	2 E72759	hypothetical prote
20	31	83.8	146	2 I40541	hypothetical prote
21	31	83.8	147	2 I39542	DNA gyrase - Aerom
22	31	83.8	147	2 AC1898	urease accessory p
23	31	83.8	148	2 F89956	conserved hypothet
24	31	83.8	155	2 S25846	homeotic protein H
25	31	83.8	156	2 T31664	DAP-kinase homolog
26	31	83.8	164	2 I58201	MHC class I antigen
27	31	83.8	171	2 T13554	hypothetical prote
28	31	83.8	178	2 T36013	probable integral
29	31	83.8	188	2 B49773	ecdysone-dependent

30	31	83.8	193	2 E83864	thioredoxin BH1717
31	31	83.8	211	2 S50686	hypothetical prote
32	31	83.8	213	2 G87457	guanylate kinase [
33	31	83.8	215	2 JX0244	pyroglutaryl-pepti
34	31	83.8	217	2 T06455	Myb26 protein - ga
35	31	83.8	223	2 I59173	glutamate decarbox
36	31	83.8	227	2 S35735	DNA gyrase chain A
37	31	83.8	228	2 T40889	hypothetical prote
38	31	83.8	228	2 S65426	pyruvate decarboxy
39	31	83.8	232	2 H82360	conserved hypothet
40	31	83.8	236	2 AB0259	probable phage min
41	31	83.8	244	2 A84729	hypothetical prote
42	31	83.8	245	2 D70882	probable dapB prot
43	31	83.8	251	2 AB0957	conserved hypothet
44	31	83.8	252	2 AE3631	nitrous-oxide redu
45	31	83.8	253	2 F84845	hypothetical prote

ALIGNMENTS

RESULT 1

S70617

homeotic protein goosecoid - fruit fly (Drosophila melanogaster)

N:Alternate names: homeobox protein goosecoid

C:Species: Drosophila melanogaster

C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004

C:Accession: S70617

R:Hahn, M.; Jaeckle, H.

EMBO J. 15, 3077-3084, 1996

A>Title: Drosophila goosecoid participates in neural development but not in body axis for

A:Reference number: S70617; MUID:96272167; PMID:8670808

A:Accession: S70617

A>Status: nucleic acid sequence not shown

A:Molecule type: mRNA

A:Residues: 1-419 <NA>

A:Cross-references: UNIPROT:P54366; EMBL:U52968; NID:G1399586; PIDN:AAB17948.1; PID:G1399

C:Genetics:

A:Gene: gsc

A:Cross-references: FlyBase:FBgn0010323

A:Map position: 2

C:Function:

A:Description: plays a role in neurogenesis in post-gastrula Drosophila embryos
A>Note: not required for gastrulation like Xenopus goosecoid; expressed most strongly in

C:Superfamily: fruit fly homeotic protein goosecoid; homeobox homology

C:Keywords: DNA binding; homeobox; nucleus; transcription regulation

F:287-343/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 37; DB 1; Length 419;

Best Local Similarity 100.0%; Pred. No. 36;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5

|||||

Db 253 HHPHG 257

RESULT 2

A60488

histidine-rich glycoprotein - bovine (fragments)

N:Alternate names: aurosette inhibition factor

C:Species: Bos primigenius taurus (cattle)

C>Date: 19-Mar-1993 #sequence_revision 23-Mar-1995 #text_change 07-Jul-1995

C:Accession: S35687; J02196; A60488

R:Sorensen, C.B.; Krogh-Pedersen, H.; Petersen, T.E.

FEBS Lett. 328, 285-290, 1993

A>Title: Determination of the disulphide bridge arrangement of bovine histidine-rich gly

A:Reference number: S35687; MUID:93351678; PMID:8348977

A:Accession: S35687

A:Molecule type: protein

A:Residues: 1-25;26-52;57-75;82-88;95-119;146-173;175-206;210-309;313-445 <SOR>

A>Note: 355-Gln and 368-Tyr were also found

R:Halkier, T.; Andersen, H.; Vestergaard, A.; Magnusson, S.

Db 411 HHPHG 415
|||||

RESULT 6
E70766
hypothetical protein Rv2082 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C>Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: E70766
R;Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: E70766
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-721 <COL>
A:Cross-references: UNIPROT:Q10690; GB:Z73966; GB:AL123456; NID:g3261577; PIDN:CAA98194.
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: Rv2082

Query Match 100.0%; Score 37; DB 2; Length 721;
Best Local Similarity 100.0%; Pred. No. 61;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 591 HHPHG 595

RESULT 7
D72660
probable aldehyde oxidoreductase APE0708 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C:Accession: D72660
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: D72660
A:Molecule type: DNA
A:Residues: 1-753 <KAW>
A:Cross-references: UNIPROT:Q9YE62; DDBJ:AP0000060; NID:g5104188; PIDN:BAAY9684.1; PID:g5
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0708
C:Superfamily: carbon monoxide dehydrogenase molybdoprotein

Query Match 100.0%; Score 37; DB 1; Length 753;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 696 HHPHG 700

RESULT 8
T13717
CRAG protein - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C>Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13717
R;Xu, X.Z.; Wes, P.D.; Chen, H.; Li, H.S.; Yu, M.; Morgan, S.; Liu, Y.; Montell, C. J. Biol. Chem. 273, 31297, 1998
A:Title: Retinal targets for calmodulin include proteins implicated in synaptic transmis

A:Reference number: Z17709; MUID:99030403; PMID:9813038
A:Accession: T13717
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-1441 <XUX>
A:Cross-references: UNIPROT:O96957; EMBL:Y17918; NID:g3893100; PIDN:CAA76938.1; PID:g3893
C:Genetics:
A:Gene: CRAG
A:Cross-references: FlyBase:FBgn0025864

Query Match 100.0%; Score 37; DB 2; Length 1441;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 1119 HHPHG 1123

RESULT 9
F97793
hypothetical protein RC0750 [imported] - Rickettsia conorii (strain Malish 7)
C:Species: Rickettsia conorii
C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
C:Accession: F97793
R;Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro Science 293, 2093-2098, 2001
A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii.
A:Reference number: A97700; MUID:21442074; PMID:11557893
A:Accession: F97793
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-76 <KUR>
A:Cross-references: UNIPROT:Q92HM1; GB:AE006914; PIDN:AAL03288.1; PID:g15619845; GSPDB:G A:Gene: RC0750

Query Match 83.8%; Score 31; DB 2; Length 76;
Best Local Similarity 80.0%; Pred. No. 65;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||

Db 7 HHPYG 11

RESULT 10
AH3162
IS21 family transposase (truncated) tnp [imported] - Agrobacterium tumefaciens (strain C C:Species: Agrobacterium tumefaciens
C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C:Accession: AH3162
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClellan, Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, ster, E.W.
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A:Reference number: AB2577; MUID:21608550; PMID:11743193
A:Accession: AH3162
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-84 <KUR>
A:Cross-references: UNIPROT:Q8UKS4; GB:AE008687; PIDN:AAL45718.1; PID:g17743448; GSPDB:G A:Experimental source: strain C58 (Dupont)
C:Genetics:
A:Gene: tnp
A:Genome: plasmid

Query Match 83.8%; Score 31; DB 2; Length 84;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
 ||||
 Db 47 HHPH 50

RESULT 11
 A49832
 DNA gyrase A - Staphylococcus epidermidis (fragment)
 C:Species: Staphylococcus epidermidis
 C>Date: 07-Apr-1994 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A49832
 R:Sreedharan, S.; Peterson, L.R.; Fisher, L.M.
 Antimicrob. Agents Chemother. 35, 2151-2154, 1991
 A:Title: Ciprofloxacin resistance in coagulase-positive and -negative staphylococci: role of DNA gyrase A
 A:Reference number: A49832; MUID:92102204; PMID:1662027
 A:Accession: A49832
 A>Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-94 <SRE>
 A:Cross-references: UNIPROT:P54112; GB:S72603; NID:G240996; PIDN:AAB20672.1; PID:G240997
 A>Note: sequence extracted from NCBI backbone (NCBI:72603, NCBI:P:72604)
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (F:1-94/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)
 Query Match 83.8%; Score 31; DB 2; Length 94;
 Best Local Similarity 80.0%; Pred. No. 81;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 5
 :||||
 Db 78 YHPHG 82

RESULT 12
 S49447
 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A - Acinetobacter baumannii (fragment)
 N:Alternate names: DNA gyrase chain A
 C:Species: Acinetobacter baumannii
 C>Date: 20-Feb-1995 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
 C:Accession: S49447
 R:Vila, J.; Ruiz, J.; Goni, P.; Marcos, M.; Jimenez de Anta, M.T.
 submitted to the EMBL Data Library, October 1994
 A:Reference number: S49447
 A:Accession: S49447
 A:Molecule type: DNA
 A:Residues: 1-95 <VIL>
 A:Cross-references: UNIPROT:Q43906; EMBL:X82165; NID:G558546; PIDN:CAA57655.1; PID:G558546
 C:Genetics:
 A:Gene: gyra
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (F:1-95/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)
 Query Match 83.8%; Score 31; DB 2; Length 95;
 Best Local Similarity 80.0%; Pred. No. 82;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 :||||
 Db 8 YHPHG 12

RESULT 13
 S54254
 DNA topoisomerase (ATP-hydrolyzing) (EC 5.99.1.3) chain A - Salmonella typhimurium (fragment)
 N:Alternate names: DNA gyrase chain A; type II DNA topoisomerase chain A
 C:Species: Salmonella typhimurium
 C>Date: 08-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 C:Accession: S54254
 R:Ruiz, J.; Castro, D.; Goni, P.; Borrego, J.J.; Vila, J.

submitted to the EMBL Data Library, April 1995
 A:Reference number: S54254
 A:Accession: S54254
 A:Molecule type: DNA
 A:Residues: 1-95 <RUI>
 A:Cross-references: UNIPROT:P37411; EMBL:X86695; NID:G799016; PIDN:CAA60388.1; PID:G799016
 C:Comment: The A chain is the target of quinolone antibiotics such as nalidixic acid and C:Genetics:
 A:Gene: gyra
 C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (F:1-95/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)
 F:51/Active site: Tyr #status Predicted
 Query Match 83.8%; Score 31; DB 2; Length 95;
 Best Local Similarity 80.0%; Pred. No. 82;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 :||||
 Db 6 YHPHG 10

RESULT 14
 S62940
 Probable membrane protein YNL028w - yeast (Saccharomyces cerevisiae)
 N:Alternate names: hypothetical protein N2758
 C:Species: Saccharomyces cerevisiae
 C>Date: 27-Apr-1996 #sequence_revision 03-May-1996 #text_change 09-Jul-2004
 C:Accession: S62940; S62950
 R:Andre, B.; Iraqi Housaini, I.; Urrestarazu, L.A.; Vissers, S.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S62920
 A:Accession: S62940
 A:Molecule type: DNA
 A:Residues: 1-105 <AND>
 A:Cross-references: UNIPROT:P53967; EMBL:Z71305; NID:G1301864; PID:G1301866; MIPS:YNL028w
 A:Experimental source: strain S288C
 R:Duesterhoeft, A.; Floeth, M.; Fritze, C.; Heuss-Neitzel, D.; Hilbert, H.; Moestl, D.
 submitted to the Protein Sequence Database, April 1996
 A:Reference number: S62944
 A:Accession: S62950
 A:Molecule type: DNA
 A:Residues: 1-105 <DUE>
 A:Cross-references: EMBL:Z71305; NID:G1301864; PID:G1301866; MIPS:YNL028w
 A:Experimental source: strain S288C
 C:Genetics:
 A:Cross-references: SGD:S0004973
 A:Map position: 14L
 C:Superfamily: Saccharomyces probable membrane protein YNL028w
 C:Keywords: transmembrane protein
 F:23-39/Domain: transmembrane #status predicted <TMM>
 Query Match 83.8%; Score 31; DB 2; Length 105;
 Best Local Similarity 100.0%; Pred. No. 90;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
 ||||
 Db 85 HHPH 88

RESULT 15
 F69760
 hypothetical protein yckD - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: F69760
 R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertero, C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Choi, A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
 Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.P.; Koetter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seron, akuechi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K. A.; Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: F69760
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-110 <KUN>
A:Cross-references: UNIPROT:P42402; GB:Z99105; GB:AL009126; NID:G2632457; PIDN:CAB12134.
A:Experimental source: strain 168
C:Genetics:
A:Gene: yckD
C:Superfamily: *Bacillus subtilis* hypothetical protein yckD

Query Match 83.8%; Score 31; DB 2; Length 110;
Best Local Similarity 100.0%; Pred. No. 95;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
DB 99 HHPH 102

RESULT 16
E82687
cytochrome O ubiquinol oxidase, subunit IV XF1387 [imported] - *Xylella fastidiosa* (strain
C:Species: *Xylella fastidiosa*
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: E82687
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen
Nature 406, 151-157, 2000
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.
A:Reference number: A82515; MUID:20365717; PMID:10910347
A:Note: for a complete list of authors see reference number A59328 below
A:Accession: E82687
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-113 <SIM>
A:Cross-references: UNIPROT:Q9PDJ2; GB:AE003970; GB:AE003849; NID:g9106392; PIDN:AAF8419
A:Experimental source: strain 9a5c
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Canargo, L.E.A.; Carraro, D.M.; Carrier, H
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.
submitted to GenBank, June 2000
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigr
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir
M.; Tuhako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z
A:Reference number: A59328
A:Contents: annotation
C:Genetics:
A:Gene: XF1387

Query Match 83.8%; Score 31; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 97;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
DB 3 HHPH 6

RESULT 17

S14983

extensin class I (clone w10-1 L) - tomato (fragment)

C:Species: *Lycopersicon esculentum* (tomato)

C:Date: 07-May-1998 #sequence_revision 15-May-1998 #text_change 17-Jul-1998

C:Accession: S14983

R:Showalter, A.M.; Zhou, J.; Rumeau, D.; Worst, S.G.; Varner, J.E.

Plant Mol. Biol. 16, 547-565, 1991

A:Title: Tomato extensin and extensin-like cDNAs: structure and expression in response to

A:Reference number: S14970; MUID:91329690; PMID:1714316

A:Accession: S14983

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-130 <SHO>

A:Cross-references: EMBL:X55694

A:Experimental source: cv. UC82B

C:Keywords: cell wall; glycoprotein; hydroxyproline

Query Match 83.8%; Score 31; DB 2; Length 130;

Best Local Similarity 100.0%; Pred. No. 1.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

DB 58 HHPH 61

RESULT 18

D72634

hypothetical protein APE1531 - *Aeropyrum pernix* (strain K1)C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: D72634

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*

A:Reference number: A74250; MUID:99310339; PMID:10382966

A:Accession: D72634

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-134 <KAW>

A:Cross-references: UNIPROT:Q9YBR8; DDBJ:AP000061; NID:g5104821; PIDN:BAA80530.1; PID:g51

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1531

C:Superfamily: *Aeropyrum pernix* hypothetical protein APE1531

Query Match 83.8%; Score 31; DB 2; Length 134;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

DB 107 HHPH 110

RESULT 19

E72759

hypothetical protein APE0066 - *Aeropyrum pernix* (strain K1)C:Species: *Aeropyrum pernix*

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: E72759

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aeropyr*

A:Reference number: A74250; MUID:99310339; PMID:10382966

A:Accession: E72759

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-136 <KAW>

A:Cross-references: UNIPROT:Q9YGS5; DDBJ:AP000058; NID:g5103388; PIDN:BAA78975.1; PID:d1

A;Experimental source: strain K1

C;Genetics:

A;Gene: APE0066

C;Superfamily: Aeropyrum pernix hypothetical protein APE0066

Query Match 83.8%; Score 31; DB 2; Length 136;

Best Local Similarity 100.0%; Pred. No. 1.2e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 132 HHPH 135

RESULT 20

I40541

hypothetical protein X - Pseudomonas solanacearum

C;Species: Pseudomonas solanacearum

C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004

C;Accession: I40541

R;Huang, J.; Carney, B.F.; Denny, T.P.; Weissinger, A.K.; Scheil, M.A.

J. Bacteriol. 177, 1259-1267, 1995

A;Title: A complex network regulates expression of eps and other virulence genes of Pseudomonas

A;Reference number: I40539; MUID:95173103; PMID:7868600

A;Accession: I40541

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-146 <RES>

A;Cross-references: UNIPROT:Q45417; EMBL:U18135; NID:G603068; PIDN:AAA66241.1; PID:G603068

Query Match

Best Local Similarity 83.8%; Score 31; DB 2; Length 146;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 142 HHPH 145

RESULT 21

I39542

DNA gyrase - Aeromonas salmonicida (fragment)

C;Species: Aeromonas salmonicida

C;Date: 19-Jul-1996 #sequence_revision 19-Jul-1996 #text_change 09-Jul-2004

C;Accession: I39542

R;Oppegard, H.; Sorum, H.

Antimicrob. Agents Chemother. 38, 2460-2464, 1994

A;Title: GyrA mutations in quinolone-resistant isolates of the fish pathogen Aeromonas salmonicida

A;Reference number: I39542; MUID:95142596; PMID:7840589

A;Accession: I39542

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: DNA

A;Residues: 1-147 <RES>

A;Cross-references: UNIPROT:P48369; GB:I42453; NID:G832830; PIDN:AAA87239.1; PID:G832831

C;Genetics:

A;Gene: gyrA

C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (fragment)

F.1-147/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (fragment)

Query Match

Best Local Similarity 83.8%; Score 31; DB 2; Length 147;

Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5

:||||

Db 45 YHPHG 49

RESULT 22

AC1898

urease accessory protein E [imported] - Nostoc sp. (strain PCC 7120)

C;Species: Nostoc sp. PCC 7120

A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 12-Jul-2004

C;Accession: AC1898

R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.

DNA Res. 8, 205-213, 2001

A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena

A;Reference number: AB1807; MUID:21595285; PMID:11759840

A;Accession: AC1898

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-147 <KUR>

A;Cross-references: UNIPROT:Q8YYW0; GB:BA000019; PIDN:BAW72690.1; PID:g17130078; GSPDB:G

A;Experimental source: strain PCC 7120

C;Genetics:

A;Gene: alr0733

C;Superfamily: urease accessory protein (nickel metallochaperone) UreE

Query Match

Best Local Similarity 83.8%; Score 31; DB 2; Length 147;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 144 HHPH 147

RESULT 23

F89956

conserved hypothetical protein SA1543 [imported] - Staphylococcus aureus (strain N315)

C;Species: Staphylococcus aureus

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004

C;Accession: F89956

R;Kuroda, M.; Ohta, T.; Uchiyama, I.; Baba, T.; Yuzawa, H.; Kobayashi, I.; Cui, L.; Oguchi, A.; Mizutani-Ui, Y.; Kobayashi, N.; Sawano, T.; Inoue, R.; Kaito, C.; Sekimizu, K.; Ima, A.; Shiba, T.; Hattori, M.; Ogasawara, N.; Hayashi, H.; Hiramoto, K.

Lancet 357, 1225-1240, 2001

A;Title: Whole genome sequencing of methicillin-resistant Staphylococcus aureus.

A;Reference number: A89758; MUID:21311952; PMID:11418146

A;Accession: F89956

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-148 <KUR>

A;Cross-references: UNIPROT:Q99TE2; GB:BA000018; PID:g13701517; PIDN:BAB42811.1; GSPDB:G

A;Experimental source: strain N315

C;Genetics:

A;Gene: SA1543

Query Match

Best Local Similarity 83.8%; Score 31; DB 2; Length 148;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

||||

Db 98 HHPH 101

RESULT 24

S25846

homeotic protein Hox A4, testicular - mouse (fragment)

N;Alternate names: homeotic protein Hox 1.4

C;Species: Mus musculus (house mouse)

C;Date: 13-Jan-1995 #sequence_revision 06-Feb-1998 #text_change 16-Aug-2004

C;Accession: S25847; S25846

R;Rubin, M.R.; Nguyen-Huu, M.C.

DNA Seq. 1, 329-334, 1991

A;Title: Murine embryonic spinal cord and adult testis Hox-1.4 cDNAs are identical 3' to 5'

A;Reference number: S25846; MUID:92190549; PMID:1686835

A;Accession: S25847

A;Molecule type: DNA

A;Residues: 1-28 <RUB1>

A;Cross-references: EMBL:X17346; NID:g51375; PID:g51376

A;Accession: S25846

A:Molecule type: mRNA
A:Residues: 27-155 <RUB2>
A:Cross-references: EMBL:X13538; NID:g51373; PIDN:CAA31889.1; PID:g51374
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F10-66/Domain: homeobox homology <HOX>

Query Match 83.8%; Score 31; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 98 HHPH 101

RESULT 25
T31664
DAP-kinase homolog COS1.3 - sea squirt (Ciona intestinalis) (fragment)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31664
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21050
A:Accession: T31664
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-156

A:Cross-references: UNIPROT:Q94426; EMBL:Z80904; NID:e1007749; PID:e274125; PIDN:CAB0258
C:Genetics:
A:Introns: 75/2; 131/3

Query Match 83.8%; Score 31; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 52 HHPH 55

RESULT 26
I58201
MHC class I antigen - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I58201
R:Lalanne, J.
Nucleic Acids Res. 10, 1039-1049, 1982
A:Title: Comparison of nucleotide sequences of mRNAs belonging to the mouse h-2 multigen
A:Reference number: I58201; MUID:82150234; PMID:6278432
A:Accession: I58201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-164 <RES>
A:Cross-references: UNIPROT:Q31156; GB:J00395; NID:g199338; PIDN:AAA39579.1; PID:g387446
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F13-78/Domain: immunoglobulin homology <IMM>

Query Match 83.8%; Score 31; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 158 HHPH 161

RESULT 27
T13554
hypothetical protein 14 - Bacillus phage phi-105
C:Species: Bacillus phage phi-105

A:Molecule type: mRNA
A:Residues: 27-155 <RUB2>
A:Cross-references: EMBL:X13538; NID:g51373; PIDN:CAA31889.1; PID:g51374
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F10-66/Domain: homeobox homology <HOX>

Query Match 83.8%; Score 31; DB 2; Length 155;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 98 HHPH 101

RESULT 25
T31664
DAP-kinase homolog COS1.3 - sea squirt (Ciona intestinalis) (fragment)
C:Species: Ciona intestinalis
C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31664
R:Bird, A.P.; Clark, V.; Jones, S.J.; Leitgeb, S.; Lennard, N.; Tweedie, S.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z21050
A:Accession: T31664
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-156

A:Cross-references: UNIPROT:Q94426; EMBL:Z80904; NID:e1007749; PID:e274125; PIDN:CAB0258
C:Genetics:
A:Introns: 75/2; 131/3

Query Match 83.8%; Score 31; DB 2; Length 156;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 52 HHPH 55

RESULT 26
I58201
MHC class I antigen - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 26-Jul-1996 #sequence_revision 26-Jul-1996 #text_change 09-Jul-2004
C:Accession: I58201
R:Lalanne, J.
Nucleic Acids Res. 10, 1039-1049, 1982
A:Title: Comparison of nucleotide sequences of mRNAs belonging to the mouse h-2 multigen
A:Reference number: I58201; MUID:82150234; PMID:6278432
A:Accession: I58201
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-164 <RES>
A:Cross-references: UNIPROT:Q31156; GB:J00395; NID:g199338; PIDN:AAA39579.1; PID:g387446
C:Superfamily: class I histocompatibility antigen; immunoglobulin homology
F13-78/Domain: immunoglobulin homology <IMM>

Query Match 83.8%; Score 31; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 158 HHPH 161

RESULT 27
T13554
hypothetical protein 14 - Bacillus phage phi-105
C:Species: Bacillus phage phi-105

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 09-Jul-2004
C:Accession: T13554
R:Kobayashi, K.; Okamura, K.; Inoue, T.; Sato, T.; Kobayashi, Y.
submitted to the EMBL Data Library, July 1998
A:Description: Complete nucleotide sequence of Bacillus subtilis phage phi-105.
A:Reference number: Z17688
A:Accession: T13554
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-171 <KOB>
A:Cross-references: UNIPROT:Q92XC1; EMBL:AB016282; PIDN:BAA36671.1
C:Superfamily: Bacillus phage phi-105 hypothetical protein 14

Query Match 83.8%; Score 31; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 23 HHPH 26

RESULT 28
T36013
probable integral membrane protein - Streptomyces coelicolor (fragment)
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T36013
R:Seeger, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, July 1999
A:Reference number: Z21574
A:Accession: T36013
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-178 <SE>
A:Cross-references: UNIPROT:Q8CK12; EMBL:AL096839; PIDN:CAB50766.1; GSPDB:GNO0070; SCOEDE
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDE:SCC22.23c

Query Match 83.8%; Score 31; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4
|||
Db 150 HHPH 153

RESULT 29
B49773
ecdysone-dependent cuticle protein EDG-84 - fruit fly (Drosophila melanogaster)
C:Species: Drosophila melanogaster
C:Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
C:Accession: B49773
R:Apple, R.T.; Fristrom, J.W.
Dev. Biol. 146, 569-582, 1991
A:Title: 20-Hydroxyecdysone is required for, and negatively regulates, transcription of I
A:Reference number: A49773; MUID:91323677; PMID:1713868
A:Accession: B49773
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-188 <APP>
A:Cross-references: UNIPROT:P27780; GB:M71249; NID:g157324; PIDN:AAA28501.1; PID:g157325
C:Genetics:
A:Gene: FlyBase:Edg84A
A:Cross-references: FlyBase:FBgn0000552

Query Match 83.8%; Score 31; DB 2; Length 188;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPH 4

Db 179 HHPH 182
||||
RESULT 30
E83864
thioredoxin BH1717 [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C:Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: E83864
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: E83864
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-193 <STO>
A:Cross-references: UNIPROT:Q9KCS7; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BA8054
A:Experimental source: strain C-125
C:Genetics:
A:Gene: BH1717
Query Match 83.8%; Score 31; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPH 4
Db 5 HHPH 8
||||
RESULT 31
S50686
hypothetical protein YER183c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C:Date: 28-May-1993 #sequence_revision 24-Feb-1995 #text_change 09-Jul-2004
C:Accession: S50686
R:Dietrich, F.S.
submitted to the EMBL Data Library, December 1994
A:Description: The sequence of S. cerevisiae cosmids 9163 and 9132.
A:Reference number: S50431
A:Accession: S50686
A:Molecule type: DNA
A:Residues: 1-211 <DIE>
A:Cross-references: UNIPROT:P40039; EMBL:U18922; NID:g603405; PID:g603424; GSPDB:GN00005
C:Genetics:
A:Gene: SGD:PAU1; MIPS:YER183c
A:Cross-references: SGD:S0000985
A:Map position: 5R
Query Match 83.8%; Score 31; DB 2; Length 211;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPH 4
Db 93 HHPH 96
||||
RESULT 32
G87457
guanylate kinase [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: G87457
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Emolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: G87457

A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-213 <STO>
A:Cross-references: UNIPROT:Q9A7N9; GB:AE005673; NID:g13423091; PIDN:AAK23659.1; GSPDB:G1
C:Genetics:
A:Gene: CC1681
C:Superfamily: guanylate kinase; guanylate kinase homology
Query Match 83.8%; Score 31; DB 2; Length 213;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPH 4
Db 5 HHPH 8
||||
RESULT 33
JX0244
pyroglutamyl-peptidase I (EC 3.4.19.3) - Bacillus amyloliquefaciens
N:Alternate names: 5-oxoprolyl-peptidase; pyroglutamyl aminopeptidase
C:Species: Bacillus amyloliquefaciens
C:Date: 10-Jun-1993 #sequence_revision 10-Jun-1993 #text_change 09-Jul-2004
C:Accession: JX0244
R:Yoshimoto, T.; Shimoda, T.; Kitazono, A.; Kabashima, T.; Ito, K.; Tsuru, D.
J. Biochem. 113, 67-73, 1993
A:Title: Pyroglutamyl peptidase gene from Bacillus amyloliquefaciens: Cloning, sequencing
A:Reference number: JX0244; MUID:93203177; PMID:8095933
A:Accession: JX0244
A:Molecule type: DNA
A:Residues: 1-215 <YOS>
A:Cross-references: UNIPROT:P46107; DDBJ:D11035; NID:g216315; PIDN:BAA01791.1; PID:g21631
C:Superfamily: pyrrolidone-carboxylate peptidase (N-terminal pyroglutamyl peptidase)
C:Keywords: hydrolase; omega peptidase

Query Match 83.8%; Score 31; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPH 4
Db 158 HHPH 161
||||
RESULT 34
T06455
Myb26 protein - garden pea
C:Species: Pisum sativum (garden pea)
C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
C:Accession: T06455
R:Uimari, A.; Strommer, J.
Plant J. 12, 1273-1284, 1997
A:Title: Myb26: A MYB-like protein of pea flowers with affinity for promoters of phenylp
A:Reference number: Z15690; MUID:98112025; PMID:9450341
A:Accession: T06455
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: mRNA
A:Residues: 1-217 <UIM>
A:Cross-references: UNIPROT:P93474; EMBL:Y11105; NID:g1841474; PIDN:CAA71992.1; PID:g1841
C:Superfamily: Arabidopsis myb-related protein 2; myb DNA-binding repeat homology
C:Keywords: DNA binding
F:11-63/Domain: myb DNA-binding repeat homology <MYB>
F:64-113/Domain: myb DNA-binding repeat homology <MYB1>
Query Match 83.8%; Score 31; DB 2; Length 217;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPH 4
Db 143 HHPH 146
||||

```

RESULT 35
I59173
Glutamate decarboxylase - rat
C:Species: Rattus norvegicus (Norway rat)
C>Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I59173
R:Bond, R.W.; Wyborski, R.J.; Gottlieb, D.I.
Proc. Natl. Acad. Sci. U.S.A. 87, 8771-8775, 1990
A>Title: Developmentally regulated expression of an exon containing a stop codon in the
A:Reference number: I59173; MUID:91062362; PMID:2247446
A:Accession: I59173
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: mRNA
A:Residues: 1-223 <RES>
A:Cross-references: UNIPROT:Q63211; GB:M38350; NID:G204231; PIDN:AAA41185.1; PID:G204232
C:Superfamily: human glutamate decarboxylase

Query Match      83.8%; Score 31; DB 2; Length 223;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 4
Db 141 HHPHG 144

RESULT 36
S35735
DNA gyrase chain A - Spiroplasma citri (fragment)
C:Species: Spiroplasma citri
C>Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 09-Jul-2004
C:Accession: S35735
R:Ye, F.; Laigret, F.; Bove, J.
submitted to the EMBL Data Library, December 1992
A:Description: Nucleotide sequence and genetic organization at the replication origin (c
A:Reference number: S35732
A:Accession: S35735
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-227 <YEF>
A:Cross-references: UNIPROT:P34030; EMBL:Z19108; NID:G49345; PIDN:CAA79524.1; PID:G49349
C:Genetic code: SGC3
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase (
C:Keywords: ATP; DNA binding
F:7-227/Domain: phage T4 DNA topoisomerase (ATP-hydrolyzing) medium chain homology (frag

Query Match      83.8%; Score 31; DB 2; Length 227;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 84 YHPHG 88

RESULT 37
T40889
Hypothetical protein SPCC1235.15 - fission yeast (Schizosaccharomyces pombe) (fragment)
C:Species: Schizosaccharomyces pombe
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000
C:Accession: T40889
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Murphy, L.; Harris, D.
submitted to the EMBL Data Library, September 1998
A:Reference number: Z21954
A:Accession: T40889
A>Status: preliminary; translated from GB/EMBL/DBBJ
A:Molecule type: DNA
A:Residues: 1-228 <WOO>
A:Cross-references: EMBL:AL031764; PIDN:CAA221119.1; GSPDB:GN00068; SPDB:SPCC1235.15
A:Experimental source: strain 972h-; cosmid c1235
C:Genetics:
A:Gene: SPDB:SPCC1235.15

```

```

A:Map position: 3
C:Superfamily: Caenorhabditis elegans hypothetical protein K07B1.4

Query Match      83.8%; Score 31; DB 2; Length 228;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 118 YHPHG 122

RESULT 38
S65426
Pyruvate decarboxylase (EC 4.1.1.1) - fava bean (fragment)
C:Species: Vicia faba (fava bean)
C>Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 12-Jul-2004
C:Accession: S65426
R:Muecke, U.; Wohlfarth, T.; Fiedler, U.; Baeumlein, H.; Ruecknagel, K.P.; Koenig, S.
Eur. J. Biochem. 237, 373-382, 1996
A>Title: Pyruvate decarboxylase from Pisum sativum. Properties, nucleotide and amino acid
A:Reference number: S65423; MUID:96215432; PMID:8647075
A:Accession: S65426
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-228 <MUE>
A:Cross-references: UNIPROT:Q7M228
C:Superfamily: thiamin pyrophosphate-binding domain homology
C:Keywords: carbon-carbon lyase; carboxy-lyase

Query Match      83.8%; Score 31; DB 2; Length 228;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 4
Db 95 HHPHG 98

RESULT 39
H82360
conserved hypothetical protein VC0127 [imported] - Vibrio cholerae (strain N16961 serogr
C:Species: Vibrio cholerae
C>Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004
C:Accession: H82360
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.
Nature 406, 477-483, 2000
A>Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.
A:Reference number: A82035; MUID:20406833; PMID:10952301
A:Accession: H82360
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-232 <HEI>
A:Cross-references: UNIPROT:Q9KVL5; GB:AE004103; GB:AE003852; NID:G96554523; PIDN:AAF93304
A:Experimental source: serogroup O1; strain N16961; biotype El Tor
C:Genetics:
A:Gene: VC0127
A:Map position: 1

Query Match      83.8%; Score 31; DB 2; Length 232;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 4
Db 27 HHPHG 30

RESULT 40
AB0259
Probable phage minor tail protein YP02123 [imported] - Yersinia pestis (strain CO92)

```

C:Species: Yersinia pestis
C:Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004
C:Accession: AB0259
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;
il, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,
Nature 413, 523-527, 2001
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.
A:Reference number: AB0001; MUID:21470413; PMID:11586360
A:Accession: AB0259
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-236 <KUR>
C:Cross-references: UNIPROT:Q8ZEN8; GB:AL590842; PIDN:CAC90934.1; PID:g15980130; GSPDB:Q
C:Gene: YPO2123
C:Superfamily: phage lambda tail assembly protein K

Query Match 83.8%; Score 31; DB 2; Length 236;
Best Local Similarity 80.0%; Pred. No. 2e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||:|
Db 205 HHPIG 209

Search completed: June 15, 2005, 14:22:47
Job time : 21.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-8

Perfect score: 37

Sequence: 1 HHPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	100.0	72	Q8CLE3	Q8cle3 yersinia pe
2	37	100.0	189	Q6E2P7	Q6e2p7 porana comm
3	37	100.0	192	Q6E2S2	Q6e2s2 ipomoea pes
4	37	100.0	198	Q9W4B3	Q9w4b3 drosophila
5	37	100.0	241	Q8TGD7	Q8tgd7 homo sapien
6	37	100.0	246	Q87OR5	Q87or5 vibrio para
7	37	100.0	273	1 CXB5 HUMAN	Q95377 homo sapien
8	37	100.0	284	Q9VAN2	Q9van2 drosophila
9	37	100.0	287	Q9W7B8	Q9w7b8 brachydanio
10	37	100.0	294	Q6SF52	Q6sf52 bacillus li
11	37	100.0	295	Q7QID0	Q7qid0 anopheles g
12	37	100.0	312	Q68DR3	Q68dr3 homo sapien
13	37	100.0	331	Q67PX1	Q67px1 symbiobacte
14	37	100.0	340	Q8DU31	Q8dj31 synechococc
15	37	100.0	348	Q9X6C8	Q9x6c8 thermus bro
16	37	100.0	349	Q8GEA7	Q8gea7 thermus sp.
17	37	100.0	349	Q9X6D3	Q9x6d3 thermus the
18	37	100.0	349	Q746I4	Q746i4 thermus the
19	37	100.0	356	Q73679	Q73679 brachydanio
20	37	100.0	356	Q98UK5	Q98uk5 brachydanio
21	37	100.0	396	1 HRC BOVIN	P33433 bos taurus
22	37	100.0	419	1 GSC DROME	P54366 drosophila
23	37	100.0	428	Q9PWC8	Q9pwc8 brachydanio
24	37	100.0	430	Q98SS0	Q98ss0 brachydanio
25	37	100.0	430	Q98SS3	Q98ss3 brachydanio
26	37	100.0	431	Q98SS2	Q98ss2 brachydanio
27	37	100.0	432	Q98SS1	Q98ss1 brachydanio
28	37	100.0	469	Q8MQS2	Q8mqs2 drosophila
29	37	100.0	491	1 SPEA BACHD	Q9k9k5 bacillus ha
30	37	100.0	509	Q9VW0	Q9vwo drosophila
31	37	100.0	510	Q9ESB2	Q9esb2 rattus norv

32	37	100.0	515	2 Q99PS7	Q99ps7 rattus norv
33	37	100.0	524	2 Q80XK4	Q80xk4 mus musculu
34	37	100.0	525	1 HRC_HUMAN	P04196 homo sapien
35	37	100.0	525	2 Q99PS5	Q99ps5 mus musculu
36	37	100.0	525	2 Q99PS6	Q99ps6 mus musculu
37	37	100.0	525	2 Q99PS8	Q99ps8 rattus norv
38	37	100.0	525	2 Q9ESB3	Q9esb3 mus musculu
39	37	100.0	526	1 HRC_RABIT	Q28540 oryctolagus
40	37	100.0	536	2 Q6YK32	Q6yK32 mus musculu
41	37	100.0	536	2 Q6YK32	Q6yK32 mus musculu
42	37	100.0	539	2 Q9NQA4	Q9nqa4 homo sapien
43	37	100.0	539	2 Q7TNS7	Q7tne7 mus musculu
44	37	100.0	539	2 Q9JHS6	Q9jhs6 rattus norv
45	37	100.0	539	2 Q9QYV9	Q9qyv9 rattus norv

ALIGNMENTS

RESULT 1

ID	Q8CLE3	PRELIMINARY;	PRT;	72 AA.
DT	01-MAR-2003 (TREMBLrel. 23, Created)			
DT	01-MAR-2003 (TREMBLrel. 23, Last sequence update)			
DT	01-MAR-2003 (TREMBLrel. 23, Last annotation update)			
DE	Hypothetical.			
GN	OrderedLocusNames=y1535;			
OS	Yersinia pestis.			
OC	Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;			
OC	Enterobacteriaceae; Yersinia.			
OX	NCBI_TaxID=632;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=KIMS / Biovar Mediaevalis;			
RX	MEDLINE=22137863; PubMed=12142430;			
RX	DOI=10.1128/JB.184.16.4601-4611.2002;			
RA	Deng W., Burland V., Plunkett G. III, Boutin A., Mayhew G.F., Liss P.,			
RA	Perna N.T., Rose D.J., Mau B., Zhou S., Schwartz D.C.,			
RA	Fetherston J.D., Lindler L.E., Brubaker R.R., Plano G.V.,			
RA	Straley S.C., McDonough K.A., Nilles M.L., Matson J.S., Blattner F.R.,			
RA	Perry R.D.;			
RT	"Genome sequence of Yersinia pestis KIM. ";			
RL	J. Bacteriol. 184:4601-4611(2002).			
DR	EMBL; AE013756; AAM85107.1; -.			
KW	Hypothetical protein.			
SQ	SEQUENCE 72 AA; 7848 MW; D2C5B430EE62B0DA CRC64;			

Query Match 100.0%; Score 37; DB 2; Length 72;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HHPHG 5
|||
Db 42 HHPHG 46

RESULT 2

ID	Q6E2P7	PRELIMINARY;	PRT;	189 AA.
AC	Q6E2P7;			
DT	25-OCT-2004 (TREMBLrel. 28, Created)			
DT	25-OCT-2004 (TREMBLrel. 28, Last sequence update)			
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE	Ribosomal protein L2 (Fragment).			
GN	Name=rpL2;			
OS	Porana commixta.			
OG	Chloroplast.			
OC	Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eucotids; asterids;			
OC	Lamiids; Solanales; Convolvulaceae; Dichondreae; Porana.			
OX	NCBI_TaxID=197434;			
RN	[1]			

```

RP SEQUENCE FROM N.A.
RA Stefanovic S., Olmstead R.G.;
RT "Testing the Phylogenetic Position of a Parasitic Plant (Cuscuta,
RT Convolvulaceae, Asteridae): Bayesian Inference and the Parametric
RT Bootstrap on Data Drawn from Three Genomes.";
RL Syst. Biol. 53:384-399(2004).
DR EMBL; AY596770; AAT69100.1; -. IEA.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0015934; C:large ribosomal subunit; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006412; P:protein biosynthesis; IEA.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR002171; Ribosomal L2-.
DR InterPro; IPR005880; Ribosomal_L2_b/o.
DR InterPro; IPR008991; Transl_SH3_like.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR Pfam; PF03947; Ribosomal_L2_C; 1.
DR TIGRFAMs; TIGR01171; rplB_bact; 1.
DR Chloroplast; Ribosomal protein.
FT NON_TER 1
FT NON_TER 189
SQ SEQUENCE 189 AA; 20451 MW; 2C81496B6779C7B5 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 181 HHPHG 185

RESULT 3
Q5E2S2 PRELIMINARY; PRT; 192 AA.
AC Q6E2S2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Ribosomal protein L2 (Fragment).
GN Name=rpl2;
OS Ipomoea pes-tigridis.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Convolvulaceae; Ipomoeae; Ipomoea.
OX NCBI_TaxID=89657;
RN [1]
RP SEQUENCE FROM N.A.
RA Stefanovic S., Olmstead R.G.;
RT "Testing the Phylogenetic Position of a Parasitic Plant (Cuscuta,
RT Convolvulaceae, Asteridae): Bayesian Inference and the Parametric
RT Bootstrap on Data Drawn from Three Genomes.";
RL Syst. Biol. 53:384-399(2004).
DR EMBL; AY596745; AAT69075.1; -.
DR GO; GO:0009507; C:chloroplast; IEA.
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0015934; C:large ribosomal subunit; IEA.
DR GO; GO:0003723; F:RNA binding; IEA.
DR GO; GO:0003735; F:structural constituent of ribosome; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR InterPro; IPR008994; Nucleic acid OB.
DR InterPro; IPR002171; Ribosomal_L2.
DR InterPro; IPR005880; Ribosomal_L2_b/o.
DR Pfam; PF00181; Ribosomal_L2; 1.
DR Pfam; PF03947; Ribosomal_L2_C; 1.
DR TIGRFAMs; TIGR01171; rplB_bact; 1.
DR Chloroplast; Ribosomal protein.
FT NON_TER 1
FT NON_TER 189
SQ SEQUENCE 189 AA; 20451 MW; 2C81496B6779C7B5 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 76;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 181 HHPHG 185

RESULT 4
Q9W4B3 PRELIMINARY; PRT; 198 AA.
AC Q9W4B3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DE 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CG15781-PA.
GN ORFNames=CG15781;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.A., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Frannkoch C., Baldwin D.,
RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Bottier P.,
RA Burtis K.C., Buseam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foslter C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.B., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milehina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclele J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Resse M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Sequence 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA MEDLINE=22426065; PubMed=12537568;
RA Celnikier S.E., Wheeler D.A., Krommiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,

```

RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleeb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.,
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence.",
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celnikier S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective.",
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celnikier S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review.",
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS003434; AAF46043.1; -;
DR IntAct; Q9W4B3; -;
DR FlyBase; FBgn0029774; CG15781.
SQ SEQUENCE 198 AA; 21401 MW; A0A6CAD90F6AA252 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 198;
Best Local Similarity 100.0%; Pred. No. 80;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 153 HHPHG 157

RESULT 5
Q8TDG7 PRELIMINARY; PRT; 241 AA.
ID Q8TDG7
AC Q8TDG7
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CLL-associated antigen KW-6 (Fragment).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Krackhardt A.M., Witzens M., Hodi F.S., Zauls A.J.,
RA Chessia M., Barrett P., Gribben J.G.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF432219; AAL99925.1; -;
FT NON_TER 1
SQ SEQUENCE 241 AA; 25884 MW; 33C30E6088012685 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 241;

Best Local Similarity 100.0%; Pred. No. 98;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 232 HHPHG 236

RESULT 6
Q87QR5 PRELIMINARY; PRT; 246 AA.
ID Q87QR5
AC Q87QR5
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein VP1084.
GN OrderedLocustNames=VP1084;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=RIMD 2210633 / Serotype O3:K6;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Najiima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.",
RL Lancet 363:743-749(2003).
DR EMBL; AP005076; BAC59347.1; -;
KW Complete proteome.
SQ SEQUENCE 246 AA; 27594 MW; 3F130DD284361E45 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 246;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 14 HHPHG 18

RESULT 7
CXB5_HUMAN STANDARD; PRT; 273 AA.
ID CXB5_HUMAN
AC O95377; Q9UPA3;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Gap junction beta-5 protein (Connexin 31.1) (Cx31.1).
GN Name=GJB5;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99057343; PubMed=9843209; DOI=10.1038/3840;
RA Richard G., Smith L.E., Bailey R.A., Itin P., Hohl D.,
RA Epstein E.H. Jr., DiGiovanna J.J., Compton J.G., Bale S.J.;
RT "Mutations in the human connexin gene GJB3 cause erythrodermatodermia
RT variabilis.",
RL Nat. Genet. 20:366-369(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Xia J.-H., Liu C.-Y., Zheng D., Pan Q., Xie W.;
RT "Molecular cloning of human connexin 31 and 31.1.",
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Pearce A.;

Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.

[4]
SEQUENCE FROM N.A.
TISSUE=Pancreas;
MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins B., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeng B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diachenko L., Marushina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
Brownstein M.J., Uedon T.B., Tomshyuk S., Carninci P., Prange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalish D.E.,
Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
-!- FUNCTION: One gap junction consists of a cluster of closely packed
pairs of transmembrane channels, the connexons, through which
materials of low MW diffuse from one cell to a neighboring cell.
-!- SUBUNIT: A connexon is composed of a hexamer of connexins.
-!- SUBCELLULAR LOCATION: Integral membrane protein.
-!- SIMILARITY: Belongs to the connexin family. Beta-type (group I)
subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AF099731; AAC95472.1; -;
EMBL; AF052693; AAD18005.1; -;
EMBL; AL121988; CAB90271.1; -;
EMBL; BC004379; AAH04379.1; -;
Genew; HGNC:4287; GJB5.
MIM; 604493; -;
GO; GO:0008544; P:epidermal differentiation; TAS.
InterPro; IPR000500; Connexin.
InterPro; IPR002270; Connexin311.
Pfam; PF00029; Connexin; 1.
PRINTS; PR00206; CONNEXIN.
SMART; SM00037; CNX; 1
PROSITE; PS00407; CONNEXINS_1; 1.
PROSITE; PS00408; CONNEXINS_2; 1.
Gap junction; Transmembrane.
FT DOMAIN 1 20 Cytoplasmic (Potential).
FT TRANSMEM 21 40 Potential.
FT DOMAIN 41 75 Extracellular (Potential).
FT TRANSMEM 76 98 Potential.
FT DOMAIN 99 126 Cytoplasmic (Potential).
FT TRANSMEM 127 149 Potential.
FT DOMAIN 150 187 Extracellular (Potential).
FT TRANSMEM 188 210 Potential.
FT DOMAIN 211 273 Cytoplasmic (Potential).
SQ SEQUENCE 273 AA; 31088 MW; C1B9F1A4AE070BB CRC64;
Query Match 100.0%; Score 37; DB 1; Length 273;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
|||||

Db 229 HHPHG 233

RESULT 8
Q9VAN2 PRELIMINARY; PRT; 284 AA.
ID Q9VAN2
AC Q9VAN2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-OCT-2002 (TREMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)
DE CG14503-PA.
ORFNames=CG14509;
GN Drosophila melanogaster (Fruit fly).
OS Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
Dodson K., Douc L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai X.,
Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
"The genome sequence of Drosophila melanogaster.";
Science 287:2185-2195 (2000).
[2]
SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
"Finishing a whole-genome shotgun: Release 3 of the Drosophila
melanogaster euchromatic genome sequence.";
Genome Biol. 3:RESEARCH0079-RESEARCH0079 (2002).
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;

RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskaas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celisner S.E.;
RT "The transposable elements of the *Drosophila melanogaster* euchromatin:
RT a genomics perspective.";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whittier E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review.";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF003768; AAF56872.2; -;
DR FlyBase; FBgn0039647; CG14509.
SQ SEQUENCE 284 AA; 31482 MW; A94AA7075D40446C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 284;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 240 HHPHG 244

RESULT 9
Q9W7B8 PRELIMINARY; PRT; 287 AA.
AC Q9W7B8;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Transcription factor Tcf3b (Fragment).
GN Name=tcf3b; Synonyms=tcf3b;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99376682; PubMed=10446273; DOI=10.1016/S0925-4773(99)00101-X;
RA Dorsky R.I., Snyder A., Cretekos C.J., Grunwald D.J., Geisler R.,
RA Haffter P., Moon R.T., Raible D.W.;
RT "Maternal and embryonic expression of zebrafish *lefi*.";
RL Mech. Dev. 86:147-150(1999).
DR EMBL; AF136456; AAD41491.1; -;
DR HSSP; P27782; 2LEF.
DR ZFIN; ZDB-GENE-991110-10; tcf3b11b.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009071; HMG-box.
DR DR Pfam; PF00505; HMG box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS0118; HMG BOX_2; 1.
PT NON TER 1
SQ SEQUENCE 287 AA; 31586 MW; 8573199746D91F0F CRC64;

Query Match 100.0%; Score 37; DB 2; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 12 HHPHG 16

RESULT 10
Q65F52 PRELIMINARY; PRT; 294 AA.
AC Q65F52;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE YuxN (Homedomain-like).
GN Name=yuxN; ORFNames=BL00748, BLi03485;
OS Bacillus licheniformis DSM 13.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=279010;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=DSM 13;
RX PubMed=15383718;
RA Veith B., Herzberg C., Steckel S., Feesche J., Maurer K.H.,
RA Ehrenreich P., Baumer S., Henne A., Liesegang H., Merkl R.,
RA Ehrenreich A., Gottschalk G.;
RT "The Complete Genome Sequence of *Bacillus licheniformis* DSM13, an
RT Organism with Great Industrial Potential.";
RL J. Mol. Microbiol. Biotechnol. 7:204-211(2004).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 14580;
RA Rey M.W., Ramaiva P., Nelson B.A., Brody-Karpin S.D., Zaretsky E.J.,
RA Tang M., de Leon A.L., Xiang H., Gusti V., Clausen I.G., Olsen P.B.,
RA Rasmussen M.D., Andersen J.T., Jorgensen P.L., Larsen T.S.,
RA Sorokin A., Bolotin A., Lapidus A., Galleron N., Ehrlich S.D.,
RA Berka R.M.;
RT "Complete genome sequence of the industrial bacterium *Bacillus*
RT *licheniformis* and comparisons with closely related *Bacillus* species.";
RL Genome Biol. 5:R77-R77(2004).
DR EMBL; AE017333; AAU42312.1; -;
DR EMBL; CP000002; AAU42943.1; -;
KW Nuclear protein.
SQ SEQUENCE 294 AA; 34659 MW; 8C74565A1ABA478C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 290 HHPHG 294

RESULT 11
Q7Q1D0 PRELIMINARY; PRT; 295 AA.
AC Q7Q1D0;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE AGCP3965 (Fragment).
GN Name=agc3965; ORFNames=ENSGG00000017877;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;

RA Anopheles Genome Sequencing Consortium;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AA001008807; EAA04166.1; -.
DR HSSP; Q99958; 1D5V.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR001766; TF_Fork_head.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF00250; Fork_head; 1.
DR PRINTS; PR00053; FORKHEAD.
DR ProDom; PD000425; TF_Fork_head; 1.
DR PROSITE; PS00657; FORK_HEAD_1; 1.
DR PROSITE; PS00658; FORK_HEAD_2; 1.
DR PROSITE; PS00659; FORK_HEAD_3; 1.
FT NON_TER 295
SQ SEQUENCE 295 AA; 32310 MW; CBB8AF12C2DA2970 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 295;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 272 HHPHG 276

RESULT 12

Q68DR3
ID Q68DR3 PRELIMINARY; PRT; 312 AA.
AC Q68DR3
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein DKFZp779H1622 (Fragment).
GN Name=DKFZp779H1622;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RG The German CDNA Consortium;
RA Koehrer K., Beyer A., Mewes H.W., Weil B., Amid C., Osanger A.,
RA Fobo G., Han M., Wiemann S.;
RL Submitted (AUG-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; CR749302; CAH18157.1; -.
KW Hypothetical protein.
FT NON_TER 1
SQ SEQUENCE 312 AA; 35124 MW; 08522797450AFALF CRC64;

Query Match 100.0%; Score 37; DB 2; Length 312;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 171 HHPHG 175

RESULT 13

Q67PX1
ID Q67PX1 PRELIMINARY; PRT; 331 AA.
AC Q67PX1
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=STH1287;

OS Symbiobacterium thermophilum.
OC Bacteria; Actinobacteria; Symbiobacterium.
OX NCBI_TaxID=2734;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IAM14863;
RA Ueda K., Yamashita A., Ishikawa J., Shimada M., Watsuji T.,
RA Morimura K., Ikeda H., Hattori M., Beppu T.;
RT "Complete genome sequence of an uncultured bacterium Symbiobacterium
RT thermophilum";
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP006840; BAD40272.1; -.
KW Hypothetical protein.
SQ SEQUENCE 331 AA; 36263 MW; 2BF2D7BD61A84880 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 331;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 149 HHPHG 153

RESULT 14

Q8DJ31
ID Q8DJ31 PRELIMINARY; PRT; 340 AA.
AC Q8DJ31
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE T11398 protein.
GN OrderedLocusNames=t11398;
OS Synecococcus elongatus (Thermosynechococcus elongatus).
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32046;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BP-1;
RX MEDLINE=2225144; PubMed=12240834;
RA Nakamura Y., Kaneko T., Sato S., Ikeuchi M., Katoh H., Sasamoto S.,
RA Watanabe A., Iriguchi M., Kawashima K., Kimura T., Kishida Y.,
RA Kiyokawa C., Kohara M., Matsumoto M., Matsuno A., Nakazaki N.,
RA Shimo S., Sugimoto M., Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the thermophilic cyanobacterium
RT Thermosynechococcus elongatus BP-1.";
RL DNA Res. 9:123-130(2002).
DR EMBL; AP005373; BAC08950.1; -.
DR InterPro; IPR002762; CbiX.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF01903; CbiX; 2.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 340 AA; 38706 MW; A9363ED422C0BADB CRC64;

Query Match 100.0%; Score 37; DB 2; Length 340;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||||
Db 317 HHPHG 321

RESULT 15

Q9X6C8
ID Q9X6C8 PRELIMINARY; PRT; 348 AA.
AC Q9X6C8
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase homolog.
GN Name=galt;

```

OS Thermus brockianus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=56956;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ITI360;
RX MEDLINE=99402735; PubMed=10473401;
RA Fridjoneson O., Watzlawick H., Gehweiler A., Rohrhirsch T., Mattes R.;
RT "Cloning of the gene encoding a novel thermostable alpha-galactosidase
from Thermus brockianus ITI360.";
RL Appl. Environ. Microbiol. 65:3955-3963(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=ITI360;
RX MEDLINE=20203878; PubMed=10741834;
RA Fridjoneson O., Watzlawick H., Mattes R.;
RT "The structure of the alpha-galactosidase gene loci in Thermus
brockianus ITI360 and Thermus thermophilus TH125.";
RL Extremophiles 4:23-33(2000).
CC -|- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
alpha-D-glucose 1-phosphate + UDP-galactose.
CC -|- PATHWAY: Galactose metabolism; second step.
CC -|- SIMILARITY: Belongs to the galactose-1-phosphate
uridylyltransferase family 1.
DR EMBL; AF135398; AAD33669.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . ; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . ; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; GalP_UDPtransf1.
DR InterPro; IPR005849; GalP_Utrnsf_N.
DR Pfam; PF02744; GalP_UDP_tr_C; 1.
DR PIRSF; PIRSF000808; GalT; 1.
DR ProDom; PD005051; GalP_Utrnsf; 2.
DR TIGRFAMs; TIGR00209; galT 1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF I; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transferase.
SQ SEQUENCE 348 AA; 39855 MW; BF6CA7BEA85DC904 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 348;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 183 HHPHG 187

RESULT 16
Q8GEA7 PRELIMINARY; PRT; 349 AA.
AC Q8GEA7;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase-like protein.
OS Thermus sp. IB-21.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=206164;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IB-21;
RX Kang S.K., Cho K.K., Ahn J.K., Choi Y.J.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -|- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
alpha-D-glucose 1-phosphate + UDP-galactose.

```

```

CC -|- PATHWAY: Galactose metabolism; second step.
CC -|- SIMILARITY: Belongs to the galactose-1-phosphate
uridylyltransferase family 1.
DR EMBL; AY130259; AAN05445.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . ; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . ; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; GalP_UDPtransf1.
DR InterPro; IPR005849; GalP_Utrnsf_N.
DR Pfam; PF01087; GalP_UDP_trnsf; 1.
DR PIRSF; PIRSF000808; GalT; 1.
DR ProDom; PD005051; GalP_Utrnsf; 2.
DR TIGRFAMs; TIGR00209; galT 1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF I; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transferase.
SQ SEQUENCE 349 AA; 39736 MW; DF8A239FA7C59607 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 183 HHPHG 187

RESULT 17
Q9X6D3 PRELIMINARY; PRT; 349 AA.
AC Q9X6D3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Galactose-1-phosphate uridylyltransferase.
GN Name=galT;
OS Thermus thermophilus.
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=274;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TH125;
RX MEDLINE=20203878; PubMed=10741834;
RA Fridjoneson O., Watzlawick H., Mattes R.;
RT "The structure of the alpha-galactosidase gene loci in Thermus
brockianus ITI360 and Thermus thermophilus TH125.";
RL Extremophiles 4:23-33(2000).
CC -|- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
alpha-D-glucose 1-phosphate + UDP-galactose.
CC -|- PATHWAY: Galactose metabolism; second step.
CC -|- SIMILARITY: Belongs to the galactose-1-phosphate
uridylyltransferase family 1.
DR EMBL; AF135399; AAD32629.1; -.
DR HSSP; P09148; 1GUQ.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . ; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . ; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; GalP_UDPtransf1.
DR InterPro; IPR005850; GalP_Utrnsf_C.
DR Pfam; PF01087; GalP_UDP_trnsf; 1.
DR Pfam; PF02744; GalP_UDP_tr_C; 1.
DR PIRSF; PIRSF000808; GalT; 1.
DR ProDom; PD005051; GalP_Utrnsf; 2.
DR TIGRFAMs; TIGR00209; galT 1; 1.
KW Carbohydrate metabolism; Galactose metabolism; Nucleotidyltransferase;
KW Transferase.

```

```
SQ SEQUENCE 349 AA; 39451 MW; 0B2284608704339A CRC64;
Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 183 HHPHG 187

RESULT 18
Q74614 PRELIMINARY; PRT; 349 AA.
AC Q74614;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Galactose-1-phosphate uridylyltransferase (EC 2.7.7.10).
GN OrderedLocusNames=TTF0071;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OG Thermidiales; Deinococcus-Thermus; Thermococci; Thermales; Thermaceae;
OC Bacteria;
OC Thermus.
OX NCBI_TaxID=262724;
RN SEQUENCE FROM N.A.
RX PubMed=15064768;
RA Henne A., Brueggemann H., Raasch C., Wierzer A., Hartsch T.,
RA Liesegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
RA Jacobi C., Starkuviene V., Schlentz S., Dencker S., Huber R.,
RA Klensk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.,
RA "The genome sequence of the extreme thermophile Thermus
RT thermophilus."
RL Nat. Biotechnol. 22:547-553(2004).
CC -!- CATALYTIC ACTIVITY: UDP-glucose + alpha-D-galactose 1-phosphate =
CC alpha-D-glucose 1-phosphate + UDP-galactose.
CC -!- PATHWAY: Galactose metabolism; second step.
CC -!- SIMILARITY: Belongs to the galactose-1-phosphate
CC uridylyltransferase family 1.
DR EMBL; AF017222; AAC82401.1; -
DR GO; GO:0008108; F:UDP-glucose-hexose-1-phosphate uridylyltran. . . ; IEA.
DR GO; GO:0003982; F:UTP-hexose-1-phosphate uridylyltransferase . . . ; IEA.
DR GO; GO:0006012; P:galactose metabolism; IEA.
DR InterPro; IPR001937; GalP_UDPtransf1.
DR InterPro; IPR011573; GalP_Utransf.
DR InterPro; IPR005850; GalP_Utransf.C.
DR InterPro; IPR005849; GalP_Utransf.N.
DR Pfam; PF01087; GalP_UDP_transf; 1.
DR Pfam; PF02744; GalP_UDP_tr.C; 1.
DR PIRSF; PIRSF00808; GalT; 1.
DR ProDom; PD005051; GalP_Utransf; 2.
DR TIGRFAMs; TIGR00209; galT.1; 1.
DR PROSITE; PS00117; GAL_P_UDP_TRANSF_I; 1.
KW Carbohydrate metabolism; Complete proteome; Galactose metabolism;
KW Nucleotide metabolism; Transference.
SQ SEQUENCE 349 AA; 39576 MW; 2045DC5BE440FA76 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 349;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 183 HHPHG 187

RESULT 19
O73679 PRELIMINARY; PRT; 356 AA.
AC O73679;
DT 01-AUG-1998 (TrEMBLrel. 07, Created)
DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
```

```
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcription factor Val.
GN Name=mafB; Synonym=valentino;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RX MEDLINE=98165393; PubMed=9425134;
RA Moens C.B., Cordes S.P., Giorgianni M.W., Barsh G.S., Kimmel C.B.;
RT "Equivalence in the genetic control of hindbrain segmentation in fish
RT and mouse."
RL Development 125:381-391(1998).
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AF066641; AAC18821.1; -.
DR HSSP; O54790; IKIV.
DR ZFIN; ZDB-GENE-980526-515; mafB.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk transcr_DNA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS0217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 356 AA; 40243 MW; 07420DB0F6CD08F1 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 356;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 219 HHPHG 223

RESULT 20
Q98UK5 PRELIMINARY; PRT; 356 AA.
AC Q98UK5;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Transcription factor MafB.
GN Name=mafB; Synonym=mafB;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN SEQUENCE FROM N.A.
RX MEDLINE=21064923; PubMed=11134968;
RA Kajihara M., Kawachi S., Kobayashi M., Ogino H., Takahashi S.,
RA Yasuda K.;
RT "Isolation, characterization, and expression analysis of zebrafish
RT large MafB."
RL J. Biochem. 129:139-146(2001).
CC -!- SIMILARITY: Belongs to the bZIP family.
DR EMBL; AB006322; BAB21102.1; -.
DR HSSP; O54790; IKIV.
DR ZFIN; ZDB-GENE-980526-515; mafB.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR InterPro; IPR008917; Euk transcr_DNA.
DR InterPro; IPR004827; TF_bZIP.
DR Pfam; PF03131; bZIP_Maf; 1.
```

```
DR SMART; SM00338; BRLZ; 1.
DR PROSITE; PS50217; BZIP; 1.
KW DNA-binding; Nuclear protein.
SQ SEQUENCE 356 AA; 40233 MW; DE4C96B62C058865 CRC64;

Query Match
Best Local Similarity 100.0%; Score 37; DB 2; Length 356;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 219 HHPHG 223

RESULT 21
HRG_BOVIN
ID HRG_BOVIN STANDARD; PRT; 396 AA.
AC P33433;
DT 01-FEB-1994 (Rel. 28, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein (Histidine-proline rich glycoprotein)
DE (HPRG) (Fragments).
GN Name=HRG;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE AND DISULFIDE BONDS.
RX MEDLINE=93351678; PubMed=8348977; DOI=10.1016/0014-5793(93)80945-Q;
RA Soerensen C.B., Krogh-Pedersen H., Petersen T.E.;
RT "Determination of the disulphide bridge arrangement of bovine
histidine-rich glycoprotein."
RL FEBS Lett. 328:285-290(1993).
CC -!- FUNCTION: The physiological function is not yet known. It binds
heme, eyes and divalent metal ions. It can inhibit rosette
formation and is known to interact with heparin, thrombospondin,
and the lysine-binding site of plasminogen. On the basis of its
homology with HMW kininogen, the His-rich region of this protein
may mediate the contact activation phase of intrinsic blood
coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
protein has many internal repeats. 12 tandem repetitions of a 5-
residue sequence (GHPHP, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.
DR InterPro; IPR000010; Prot_inh_cystat.
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat.
FT DOMAIN 1 102 Cystatin-like 1.
FT DOMAIN 103 169 Cystatin-like 2.
FT DOMAIN 191 238 Pro-rich.
FT DOMAIN 243 368 His/Pro-rich.
FT DISULFID 7 375
FT DISULFID 56 67
FT DISULFID 77 92
FT DISULFID 123 297
FT DISULFID 137 160
FT DISULFID 212 242
FT NON_CONS 52 53
FT CARBOHYD 70 70 N-linked (GlcNAc. . .).
FT NON_CONS 71 72
FT NON_CONS 78 79
FT CARBOHYD 91 91 N-linked (GlcNAc. . .).
FT NON_CONS 103 104
FT CARBOHYD 122 122 N-linked (GlcNAc. . .).
FT NON_CONS 163 164
FT CARBOHYD 220 220 N-linked (GlcNAc. . .).
FT NON_CONS 263 264
FT NON_CONS 303 304
FT VARIANT 86 S -> R.

RESULT 22
GSC_DROME
ID GSC_DROME STANDARD; PRT; 419 AA.
AC P54366; Q9VPR9;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 25-JAN-2005 (Rel. 46, Last annotation update)
DE Homeobox protein goosecooid.
GN Name=Gsc; ORFNames=CG2851;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=96202483; PubMed=8625850;
RA Coriely A., Stella M., Coffinier C., Kessler D., Mailhos C.,
RA Dessein S., Desplan C.;
RT "A functional homolog of goosecooid in Drosophila."
RL Development 122:1641-1650(1996).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96272167; PubMed=8670808;
RA Hahn M., Jackie H.;
RT "Drosophila goosecooid participates in neural development but not in
body axis formation."
RN EMBO J. 15:3077-3084(1996).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.K., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostier C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong P., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusken D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
```

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhou Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*";
RL Science 287:2185-2195(2000).
RN [4]
RP GENOME REANNOTATION.
RX MEDLINE=22426069; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochuk S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celisner S.E., de Grey A.D.N.J., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22(2002).
CC -!- FUNCTION: Appears to regulate regional development of specific
CC tissues. Can rescue axis polarity in UV-irradiated *Xenopus* embryos.
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: In early embryo development, expression
CC confined to two regions; a horseshoe-like pattern across the
CC dorsal side which is destined to form the brain hemispheres and a
CC second domain which invaginates inside the stomodaeum and which, is
CC fated to form the foregut, ring gland and stomatogastric nervous
CC system (SNS).
CC -!- SIMILARITY: Belongs to the paired homeobox family. Bicoid
CC subfamily.
CC -!- SIMILARITY: Contains 1 homeobox domain.
CC -!- CAUTION: It is uncertain whether Met-1 or Met-5 is the initiator.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/annouce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X95420; C2A64699.1; -;
DR EMBL; U52968; AAB17948.1; -;
DR EMBL; AE003589; AAF51473.1; -;
DR PIR; S70617; S70617.
DR HSP; P06601; 1FJL.
DR IntAct; P54366; -;
DR TRANSFAC; T04041; -;
DR FlyBase; FBgn0010323; Gsc.
DR InterPro; IPR001356; Homeobox.
DR InterPro; IPR009057; Homeobox domain like.
DR InterPro; IPR007104; Paired_homeo.
DR Pfam; PF00046; Homeobox; 1.
DR PRINTS; PR00024; HOMEBOX.
DR ProDom; PD000010; Homeobox; 1.
DR SMART; SM00389; HOX; 1.
DR PROSITE; PS00027; HOMEBOX 1; 1.
DR PROSITE; PS00071; HOMEBOX 2; 1.
KW Developmental protein; DNA-binding; Homeobox; Nuclear protein.
FT DOMAIN 104 107 Poly-Ala.
FT DOMAIN 164 169 Poly-Ser.
FT DOMAIN 195 199 Poly-Ala.
FT DNA_BIND 286 345 Homeobox.
SQ SEQUENCE 419 AA; 44949 MW; 851A4C46AA861FB9 CRC64;
Query Match 100.0%; Score 37; DB 1; Length 419;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 253 HHPHG 257
RESULT 23
Q9PWC8 PRELIMINARY; PRT; 428 AA.
ID Q9PWC8
AC Q9PWC8
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2004 (TrEMBLrel. 13, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2.
GN Name=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=93326148; PubMed=10395930; DOI=10.1016/S0167-4781(99)00066-4;
RA Wang H., Gong Z.;
RT "Characterization of two zebrafish cDNA clones encoding egg envelope
RT proteins ZP2 and ZP3";
RL Biochim. Biophys. Acta 1446:156-160(1999).
DR EMBL; AF095456; AAD49112.1; -;
DR ZFIN; ZDB-GENE-991129-5; ZP2.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil; 1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELUCIDA.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Envelope protein.
SQ SEQUENCE 428 AA; 46804 MW; 46E8825ACA17C6A7 CRC64;
Query Match 100.0%; Score 37; DB 2; Length 428;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HHPHG 5
Db 37 HHPHG 41
RESULT 24
Q98SS0 PRELIMINARY; PRT; 430 AA.
ID Q98SS0
AC Q98SS0
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2.
GN Name=ZP2.2; Synonyms=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21017556; PubMed=11144219;
RX DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;
RA Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;
RT "Cluster of genes encoding the major egg envelope protein of
RT zebrafish";
RL Mol. Reprod. Dev. 58:4-14(2001).
DR EMBL; AF331968; AAK16580.1; -;

```
DR ZFIN; ZDB-GENE-010622-1; zp2.2.
DR GO; GO:0019031; C:Viral envelope; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil; 1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Envelope protein.
SQ SEQUENCE 430 AA; 47124 MW; EC544EC03A10B63 CRC64;

Query Match      100.0%; Score 37; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      40 HHPHG 44

RESULT 25
Q98SS3
ID Q98SS3 PRELIMINARY; PRT; 430 AA.
AC Q98SS3
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2 variant A.
GN Name=ZP2.2; Synonyms=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21017556; PubMed=11144219;
DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;
Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;
"Cluster of genes encoding the major egg envelope protein of
zebrafish.";
RL Mol. Reprod. Dev. 58:4-14(2001).
DR EMBL; AF331966; AAK16577.1; -.
DR ZFIN; ZDB-GENE-010622-2; zp2.3.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil; 1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Envelope protein.
SQ SEQUENCE 431 AA; 47233 MW; 0529736B8A944358 CRC64;

Query Match      100.0%; Score 37; DB 2; Length 431;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      40 HHPHG 44

RESULT 27
Q98SS1
ID Q98SS1 PRELIMINARY; PRT; 432 AA.
AC Q98SS1
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Egg envelope protein ZP2 variant C.
GN Name=ZP2.4; Synonyms=ZP2;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21017556; PubMed=11144219;
DOI=10.1002/1098-2795(200101)58:1<4::AID-MRD2>3.0.CO;2-P;
Mold D.E., Kim I.F., Tsai C.-M., Lee D., Chang C.-Y., Huang R.C.;
"Cluster of genes encoding the major egg envelope protein of
zebrafish.";
RL Mol. Reprod. Dev. 58:4-14(2001).
DR EMBL; AF331967; AAK16579.1; -.
DR ZFIN; ZDB-GENE-010622-3; zp2.4.
DR GO; GO:0019031; C:viral envelope; IEA.
DR InterPro; IPR001507; Endoglin/CD105.
DR InterPro; IPR000519; P_trefoil.
DR Pfam; PF00088; Trefoil; 1.
DR Pfam; PF00100; Zona pellucida; 1.
DR PRINTS; PR00023; ZPELLUCIDA.
DR SMART; SM00018; PD; 1.
DR SMART; SM00241; ZP; 1.
DR PROSITE; PS00682; ZP_DOMAIN; 1.
KW Envelope protein.

Query Match      100.0%; Score 37; DB 2; Length 430;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
DB      40 HHPHG 44

RESULT 26
Q98SS2
ID Q98SS2 PRELIMINARY; PRT; 431 AA.
AC Q98SS2
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
```

```
SQ SEQUENCE 432 AA; 47343 MW; B985118D36D5B8C CRC64;
Query Match 100.0%; Score 37; DB 2; Length 432;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 40 HHPHG 44

RESULT 28
Q8MQS2
ID Q8MQS2 PRELIMINARY; PRT; 469 AA.
AC Q8MQS2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE GH11945P.
GN ORFNames=CG14509;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
RA Champe M., Chavez C., Dorsett V., Dresnek D., Farfan D., Frise E.,
RA George R., Gonzalez M., Guarin H., Kronmiller B., Li P., Liao G.,
RA Miranda A., Mungall C.J., Nunoo J., Pacleb J., Paragas V., Park S.,
RA Patel S., Phouanavong S., Wan K., Yu C., Lewis S.E., Rubin G.M.,
RA Celniker S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY128418; AAW75011.1; -.
DR FlyBase; FBgn0039647; CG14509.
SQ SEQUENCE 469 AA; 51845 MW; BECE1D5C4D18CC06 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 469;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 425 HHPHG 429

RESULT 29
SPEA_BACHD
ID SPEA_BACHD STANDARD; PRT; 491 AA.
AC Q9K9K5;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Arginine decarboxylase (EC 4.1.1.19).
GN Name=speA; OrderedLocustNames=BH2640;
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogawara N., Kuhara S.,
RA Horikoshi K.;
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -!- FUNCTION: Catalyzes the formation of agmatine from arginine (By
similarity).
CC -!- CATALYTIC ACTIVITY: L-arginine = agmatine + CO(2).
```

```
CC -!- COPACTOR: Pyridoxal phosphate (By similarity).
CC -!- PATHWAY: Putrescine biosynthesis from arginine; first step.
CC -!- PATHWAY: Spermidine biosynthesis from arginine; first step.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -!- SIMILARITY: Belongs to the Orn/Lys/Arg decarboxylase class-I
family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP001516; BAB06359.1; -.
DR PIR; H83979; H83979.
DR InterPro; IPR000310; Decarboxylase1.
DR InterPro; IPR008286; Decarboxylase_C.
DR InterPro; IPR011185; Lys decarb.
DR Pfam; PF01276; OKR_DC_1_1.
DR Pfam; PF03711; OKR_DC_1_C; 1.
DR PIRSF; PIRSF005938; Lys decarb; 1.
DR PROSITE; PS00703; OKR_DC_1; 1.
KW Complete proteome; Decarboxylase; Lyase; Polyamine biosynthesis;
KW Putrescine biosynthesis; Pyridoxal phosphate; Spermidine biosynthesis.
FT BINDING 227 227 Pyridoxal phosphate (By similarity).
SQ SEQUENCE 491 AA; 53449 MW; 3C2ACC55C357F5D CRC64;

Query Match 100.0%; Score 37; DB 1; Length 491;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 62 HHPHG 66

RESULT 30
QSVIWO
ID QSVIWO PRELIMINARY; PRT; 509 AA.
AC QSVIWO;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE CG10034-PA (Traffic jam).
GN ORFNames=CG10034;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yeandle M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abrial J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.V., Benos P.V., Bereman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K.J., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fowler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
```

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levcintosh A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M., M.G.,
RA Palazolo M., Pittman K.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington G., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yen R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster";
RL Science 287:2185-2195(2000).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426065; PubMed=12537568;
RX Celnikier S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirkas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence";
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RN SEQUENCE FROM N.A.
RP MEDLINE=22426070; PubMed=12537573;
RX Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirkas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celnikier S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:
RT a genomics perspective";
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RN SEQUENCE FROM N.A.
RX MEDLINE=22426089; PubMed=12537572;
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradscky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celnikier S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review";
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RN SEQUENCE FROM N.A.
RP FlyBase;
RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RN SEQUENCE FROM N.A.
RP FlyBase;
RG Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RN [7]
RN SEQUENCE FROM N.A.
RX MEDLINE=22954636; PubMed=14578908; DOI=10.1038/ncb1058;
RA Li M.A., Alls J.D., Avancini R.M., Koo K., Godt D.;
RT "The large Maf factor Traffic Jam controls gonad morphogenesis in
RT Drosophila";
RL Nat. Cell Biol. 5:994-1000(2003).
DR EMBL; AB003663; AAF53804.2; -;
DR EMBL; AY325814; AAP88969.1; -;
DR HSSP; O54790; 1K1V.
DR IntAct; Q9V1W0; -;

DR FlyBase; FBgn0000964; tj.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR008917; Euk transcr_DNA.
DR InterPro; IPR004827; TF_ZIP.
DR InterPro; IPR004826; TF_Maf.
DR Pfam; PF03131; bZIP_Maf; 1.
DR SMART; SM00338; BRLZ; 1.
SQ SEQUENCE 509 AA; 54043 MW; 0BD429AFB0BCCF38 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 267 HHPHG 271

RESULT 31
Q9ESB2 PRELIMINARY; PRT; 510 AA.
ID Q9ESB2
AC Q9ESB2;
DT 01-MAR-2001 (TREMELrel. 16, Created)
DT 01-MAR-2001 (TREMELrel. 16, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name:Hrg;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Lewis;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Farish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
RT cellular origin";
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194029; AAG28417.1; -;
DR GO; GO:0004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cyscat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 510 AA; 57581 MW; 508E6E06AA2ED58E CRC64;

Query Match 100.0%; Score 37; DB 2; Length 510;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
DB 346 HHPHG 350

RESULT 32
Q99PS7 PRELIMINARY; PRT; 515 AA.
ID Q99PS7
AC Q99PS7;
DT 01-JUN-2001 (TREMELrel. 17, Created)
DT 01-JUN-2001 (TREMELrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMELrel. 24, Last annotation update)
DE Histidine-rich glycoprotein 2.
GN Name:RNHRG2;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus;
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Sprague-Dawley; TISSUE=Liver;

```
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055896; BAB33093.1; -.
DR GO; GO:0004869; F.cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; 1.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 515 AA; 58055 MW; 7CEBA3A1A3678966 CRC64;

Query Match      100.0%; Score 37; DB 2; Length 515;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      346 HHPHG 350

RESULT 33
Q80XK4 PRELIMINARY; PRT; 524 AA.
AC Q80XK4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Accn4 protein (Fragment).
GN Name=Accn4;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skaleka U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the amiloride-sensitive sodium channel
CC family.
DR EMBL; BC046481; AAH46481.1; -.
DR MGD; MGI:2652846; Accn4.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005272; F:sodium channel activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0006814; P:sodium ion transport; IEA.
DR InterPro; IPR001873; Na-channel_ASC.
DR Pfam; PF00858; ASC; 1.

DR PRINTS; PR01078; AMINACHANNEL.
KW Ion transport; Ionic channel; Sodium channel; Sodium transport;
KW Transmembrane; Transport.
FT NON_TER 1
SQ SEQUENCE 524 AA; 57469 MW; AD1B1991F9B82FAF CRC64;

Query Match      100.0%; Score 37; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HHPHG 5
Db      509 HHPHG 513

RESULT 34
HRG_HUMAN STANDARD; PRT; 525 AA.
ID _HRG_HUMAN
AC P04196;
DT 20-MAR-1987 (Rel. 04, Created)
DT 20-MAR-1987 (Rel. 04, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein precursor (Histidine-proline rich
DE Glycoprotein) (HPRG).
GN Name=HRG;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Liver;
RX MEDLINE=86216149; PubMed=3011081;
RA Koide T., Foster D.C., Yoshitake S., Davie E.W.;
RT "amino acid sequence of human histidine-rich glycoprotein derived from
RL Biochemistry 25:2220-2225(1986).
RN [2]
RP SEQUENCE FROM N.A.
RA Wakabayashi S., Takahashi K., Tokunaga F., Koide T.;
RL Submitted (JUL-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 214-247 FROM N.A.
RX MEDLINE=94245171; PubMed=8188234;
RA Hennis B.C., Frants R.R., Bakker E., Vossen R.H., van der Poort E.W.,
RA Blondin L.A., Cox S., Khan P.M., Spurr N.K., Kluff C.;
RT "Evidence for the absence of intron H of the histidine-rich
RT glycoprotein (HRG) gene: genetic mapping and in situ localization of
RL HRG to chromosome 3q28-q29.";
RN Genomics 19:195-197(1994).
RN [4]
RP SEQUENCE OF 19-27.
RC TISSUE=Plasma;
RX MEDLINE=93093937; PubMed=1459097;
RA Hughes G.J., Frutiger S., Paquet N., Ravier F., Pasquali C.,
RA Sanchez J.-C., James R., Tissot J.-D., Bjellqvist B.,
RA Hochstrasser D.F.;
RT "Plasma protein map: an update by microsequencing.";
RL Electrophoresis 13:707-714(1992)
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats, 12 tandem repetitions of a 5-
CC residue sequence (GHHPH, consensus) form a histidine-rich region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.
-----
```

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

DR EMBL; M13149; AAA52694.1; -;
 DR EMBL; AB005803; BAA21613.1; -;
 DR EMBL; Z17218; CAA78925.1; -;
 DR PIR; A01287; KGHUGH.
 DR SWISS-2DPAGE; P04196; HUMAN.
 DR Genew; HGNC:5181; HRG.
 DR MIM; 142640; -;
 DR InterPro; IPR000010; Prot_inh_cystat.
 DR Pfam; PF00031; Cystatin; 1.
 KW Direct protein sequencing; Glycoprotein; Heparin-binding;
 KW Polymorphism; Repeat; Signal.
 FT SIGNAL 1 18
 FT CHAIN 19 525 Histidine-rich glycoprotein.
 FT DOMAIN 19 136 Cystatin-like 1.
 FT DOMAIN 137 254 Cystatin-like 2.
 FT DOMAIN 276 321 Pro-rich.
 FT DOMAIN 350 497 His/Pro-rich.
 FT DISULFID 24 504 By similarity.
 FT DISULFID 78 89 By similarity.
 FT DISULFID 105 126 By similarity.
 FT DISULFID 203 417 By similarity.
 FT DISULFID 218 241 By similarity.
 FT CARBOHYD 63 63 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 87 87 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 125 125 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 344 344 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 345 345 N-linked (GlcNAc...) (Potential).
 FT VARIANT 204 204 P -> S (in dbSNP:3181917).
 FT /FTID=VAR_014528.
 SQ SEQUENCE 525 AA; 59578 MW; A2B124D6CE93114F CRC64;

Query Match 100.0%; Score 37; DB 1; Length 525;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 384 HHPHG 388

RESULT 35
 Q99PS5 PRELIMINARY; PRT; 525 AA.
 AC Q99PS5;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
 DE Histidine-rich glycoprotein.
 GN Name=Hrg; Synonyms=MMHRG;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Tauchida N., Wakabayashi S., Jannen-Dechent W., Koide T.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 [2]
 RC STRAIN=FVB/N; TISSUE=Liver;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.W., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S.K., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=FVB/N; TISSUE=Liver;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB055898; BAB33095.1; -;
 DR EMBL; BC011168; AAH11168.1; -;
 DR MGD; MGI:2146636; Hrg.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR InterPro; IPR000010; Prot_inh_cystat.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 2.
 SQ SEQUENCE 525 AA; 59162 MW; A83E93A439CFB126 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 361 HHPHG 365

RESULT 36
 Q99PS6 PRELIMINARY; PRT; 525 AA.
 AC Q99PS6;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Histidine-rich glycoprotein.
 GN Name=Hrg; Synonyms=MMHRG;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Liver;
 RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
 RA Ohashi T., Sato N., Hirata D., Teuchida N., Koide T.;
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AB055897; BAB33094.1; -;
 DR MGD; MGI:2146636; Hrg.
 DR GO; GO:0005615; C:extracellular space; TAS.
 DR InterPro; IPR000010; Prot_inh_cystat.
 DR Pfam; PF00031; Cystatin; 1.
 DR SMART; SM00043; CY; 2.
 SQ SEQUENCE 525 AA; 59050 MW; A83E93A439CFB3AC CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
 |||||
 Db 361 HHPHG 365

```

RESULT 37
Q99PS8
ID Q99PS8 PRELIMINARY; PRT; 525 AA.
AC Q99PS8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Histidine-rich glycoprotein 1.
DE Histidine-rich glycoprotein 1.
GN Name=RNHRG1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Liver;
RA Wakabayashi S., Takahashi K., Hirokado Y., Togo Y., Izumi S.,
RA Ohashi T., Sato N., Hirata D., Tsuchida N., Koide T.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB055895; BAB33092.1; -.
DR GO; GO:0004869; F-cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59049 MW; 38290AG31FAC7777 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 346 HHPHG 350

RESULT 38
Q9ESB3
ID Q9ESB3 PRELIMINARY; PRT; 525 AA.
AC Q9ESB3;
DT 01-WAR-2001 (TrEMBLrel. 16, Created)
DT 01-WAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Histidine-rich glycoprotein.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129;
RX MEDLINE=20307726; PubMed=10849117;
RA Hulett M.D., Farish C.R.;
RT "Murine histidine-rich glycoprotein: cloning, characterization and
RT cellular origin.";
RL Immunol. Cell Biol. 78:280-287(2000).
DR EMBL; AF194028; AAG28416.1; -.
DR MGD; MGI:2146636; Hrg.
DR GO; GO:0005615; C:extracellular space; TAS.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 525 AA; 59132 MW; 6E55F2A439CFB123 CRC64;

Query Match 100.0%; Score 37; DB 2; Length 525;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
Db 361 HHPHG 365

```

```

RESULT 39
HRG_RABIT
ID HRG_RABIT STANDARD; PRT; 526 AA.
AC Q26640;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Histidine-rich glycoprotein precursor (Histidine-proline rich
DE glycoprotein) (HPRG) (Fragment).
GN Name=HRG;
OS Oryctolagus cuniculus (Rabbit).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.
OX NCBI_TaxID=9986;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 9-23; 301-313 AND 422-429.
RC TISSUE=Serum;
RX MEDLINE=96229917; PubMed=8639676; DOI=10.1021/bi952061t;
RA Borza D.-B., Tatum F.M., Morgan W.T.;
RT "Domain structure and conformation of histidine-proline-rich
RT glycoprotein.";
RL Biochemistry 35:1925-1934(1996).
CC -!- FUNCTION: The physiological function is not yet known. It binds
CC heme, dyes and divalent metal ions. It can inhibit rosette
CC formation and is known to interact with heparin, thrombospondin,
CC and the lysine-binding site of plasminogen. On the basis of its
CC homology with HMW kininogen, the His-rich region of this protein
CC may mediate the contact activation phase of intrinsic blood
CC coagulation cascade.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Expressed by the liver and secreted in plasma.
CC -!- DOMAIN: In addition to having a high His and Pro content, this
CC protein has many internal repeats. 15 tandem repetitions of a 5-
CC residue sequence (G[H/P][H/P]PH, consensus) form a His/Pro-rich
CC region.
CC -!- SIMILARITY: Contains 2 cystatin-like domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; U32189; AAC48516.1; -.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
KW Direct protein sequencing; Glycoprotein; Heparin-binding; Repeat;
KW Signal.
FT NON_TER 1
FT SIGNAL <1 8 Potential.
FT CHAIN 9 526 Histidine-rich glycoprotein.
FT DOMAIN 9 126 Cystatin-like 1.
FT DOMAIN 127 243 Cystatin-like 2.
FT DOMAIN 251 296 Pro-rich.
FT DOMAIN 329 498 His/Pro-rich.
FT DISULFID 14 505 By similarity.
FT DISULFID 68 79 By similarity.
FT DISULFID 95 116 By similarity.
FT DISULFID 193 415 By similarity.
FT DISULFID 207 230 By similarity.
FT DISULFID 272 302 Potential.
FT CARBOHYD 115 115 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 192 192 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 240 240 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 310 310 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 485 485 N-linked (GlcNAc...) (Potential).
FT SITE 303 304 Cleavage (by plasmin).
FT SITE 421 422 Cleavage (by plasmin).
SQ SEQUENCE 526 AA; 58877 MW; 810F23D367D93D42 CRC64;

```

Query Match 100.0%; Score 37; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 333 HHPHG 337

RESULT 40

Q6YK32
ID O6YK32 PRELIMINARY; PRT; 536 AA.
AC O6YK32;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Histidine-rich glycoprotein HRG.
GN Name=Hrg;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Heu S.J., Balmain A.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY137504; AAN27996.1; -.
DR GO; GO:004869; F:cysteine protease inhibitor activity; IEA.
DR InterPro; IPR000010; Prot_inh_cystat.
DR Pfam; PF00031; Cystatin; I.
DR SMART; SM00043; CY; 2.
SQ SEQUENCE 536 AA; 60439 MW; CCF14FC08DD53D7C CRC64;

Query Match 100.0%; Score 37; DB 2; Length 536;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HHPHG 5
|||
Db 372 HHPHG 376

Search completed: June 15, 2005, 14:21:24
Job time : 85.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A Genesepc16Dec04:*

- 1: Genesepc19808:*
- 2: Genesepc19908:*
- 3: Genesepc20008:*
- 4: Genesepc20018:*
- 5: Genesepc20028:*
- 6: Genesepc20038:*
- 7: Genesepc20038a:*
- 8: Genesepc20048:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	5	5	Abb79810 Histidine
2	36	100.0	5	8	Adh10417 Rabbit HP
3	36	100.0	20	4	Abb42309 Peptide #
4	36	100.0	20	4	Aam36117 Peptide #
5	36	100.0	20	4	Abb25810 Protein #
6	36	100.0	20	4	Aam76009 Human bon
7	36	100.0	20	4	Aam63195 Human bra
8	36	100.0	20	4	Abg57735 Human liv
9	36	100.0	20	5	Abg45414 Human pep
10	36	100.0	53	4	Aau46033 Propionib
11	36	100.0	53	6	Abm42552 Propionib
12	36	100.0	61	4	Aau54374 Propionib
13	36	100.0	61	6	Abm50893 Propionib
14	36	100.0	63	4	Aau22059 Human car
15	36	100.0	63	7	Ade46027 Human car
16	36	100.0	63	8	Adj07445 Human car
17	36	100.0	87	4	Aau66176 Propionib
18	36	100.0	87	6	Abm62695 Propionib
19	36	100.0	92	4	Aau58530 Propionib
20	36	100.0	92	6	Abm55049 Propionib
21	36	100.0	94	4	Aau63724 Propionib
22	36	100.0	94	6	Abm60243 Propionib
23	36	100.0	99	4	Aau60014 Propionib
24	36	100.0	99	4	Abb15846 Human ner
25	36	100.0	99	6	Abm56533 Propionib

26	36	100.0	101	8	ADH10412	Adh10412 Rabbit HP
27	36	100.0	112	4	AAO13856	AAO13856 Human pol
28	36	100.0	117	5	ABG70306	ABG70306 Human MDD
29	36	100.0	118	4	AAU00910	Aau00910 Human can
30	36	100.0	202	4	AAU40328	Aau40328 Propionib
31	36	100.0	202	6	ABM36847	Abm36847 Propionib
32	36	100.0	231	2	AAV29941	Aay29941 Zea mays
33	36	100.0	241	6	ADA48350	Ada48350 Rice prot
34	36	100.0	262	4	AAU87400	Aau87400 Novel cen
35	36	100.0	262	7	ABO61857	Abos1857 Kiebsatell
36	36	100.0	262	8	ADI54715	Adi54715 Novel hum
37	36	100.0	311	4	ABB66594	Abb66594 Drosophil
38	36	100.0	324	4	AAU51325	Aau51225 Propionib
39	36	100.0	324	6	ABM47744	Abm47744 Propionib
40	36	100.0	475	6	ABU33219	Abu33219 Protein e
41	36	100.0	480	7	ABO78053	Abos78053 Pseudomon
42	36	100.0	504	4	ABB68243	Abb68243 Drosophil
43	36	100.0	516	7	ABO71042	Abos71042 Pseudomon
44	36	100.0	526	5	ABB79805	Abb79805 Rabbit hi
45	36	100.0	526	8	ADH10410	Adh10410 Rabbit HP

ALIGNMENTS

RESULT 1

ABB79810
ID ABB79810 standard; peptide; 5 AA.

XX ABB79810;

DT 25-NOV-2002 (first entry)

DE Histidine proline rich glycoprotein pentapeptide.

KW Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW neurotropic; neuroprotective; antiparkinsonian.

OS Synthetic.

PN WO200264621-A2.

XX 22-AUG-2002.

PF 14-FEB-2002; 2002WO-US004336.

XX 14-FEB-2001; 2001US-0268370P.

PA (ATTE-) ATTENUON LLC.

PI Donate F, Harris S, Plunkett ML, Mazar AP;

XX WPI; 2002-666989/71.

DR New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
XX diagnosing or treating diseases associated with undesired cell migration,
PT invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.

XX Claim 2; Page 67; 82pp; English.

PS The present sequence is a specific example of claimed anti-angiogenic
XX pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC peptides comprise: the histidine-proline-rich (H/P) domain of human
CC histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC of rabbit HPRG (see ABB79807); a variant of these that is capable of
CC inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC tube formation in vitro or in vivo; or a pentapeptide having the generic
CC sequence given in ABB79808, such as the present peptide, or its variant
CC having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC N- or C-terminus. Also claimed are: chemically synthesised or
CC recombinantly produced peptide multimers; a diagnostically or

therapeutically labeled anti-angiogenic polypeptide, peptide or peptide multimer; a diagnostically useful HPRG-related composition, comprising the diagnostically labeled polypeptide, peptide or peptide multimer and a carrier; an antibody specific for an epitope of HPRG that is present in the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any of the domains in a way which inhibits the anti-angiogenic activity of HPRG or the domain, or an antigen-binding fragment of the antibody; a method for inhibiting cell migration, cell invasion, cell proliferation or angiogenesis, or for inducing apoptosis; a method for treating a subject having a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis; a method of stimulating or inhibiting angiogenesis in a subject; a method of detecting the presence of HPRG or its cleavage product or its peptide in a biological sample; isolated nucleic acids encoding the polypeptide, peptide or peptide multimer; an expression vector; transfected or transfecting cells; a method of providing to a cell, tissue or organ an angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer; an affinity ligand useful for binding to, or isolating, an HPRG-binding molecule or cells expressing the binding molecule, comprising the polypeptide, peptide or peptide multimer, immobilised to a solid support or carrier; and a method of isolating HPRG-binding molecule, or isolating or enriching cells expressing HPRG-binding site or receptor, from a complex mixture. The compositions and methods are useful in diagnosing or treating a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or fractures. HPRG is especially useful in inhibiting the growth of primary tumours or metastases, and may also be used in treating neurodegenerative diseases like Alzheimer's or Parkinson's disease. The antibodies are stimulators of angiogenesis and are useful for promoting CC neovascularization in pertinent disease states, and in various CC immunoassays

XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 36; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1 HPPHG 5

RESULT 2
ADH10417
XX ADH10417;
AC ADH10417;
XX 11-MAR-2004 (first entry)

DT Rabbit HPRG protein H/P rich domain repeat fragment.
XX
DE

XX Tpm; tropomyosin; antiangiogenic receptor;
KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
KW ophthalmological; antiinflammatory; gynaecological; antiarthritis;
KW antipsoriatic; dermatological; cardiac; vasotropic; vulnary;
KW angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

OS WO2003077872-A2.

PN 25-SEP-2003.

PD 17-MAR-2003; 2003WO-US008060.

PF 15-MAR-2002; 2002US-0364047P.

PR (ATTE-) ATTENUON LLC.

XX

PI Mcrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX Example 5; SEQ ID NO 31; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX
SQ Sequence 5 AA;

Query Match 100.0%; Score 36; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1 HPPHG 5

RESULT 3
ABB42309
ID ABB42309 standard; peptide; 20 AA.

AC ABB42309;

XX 04-FEB-2002 (first entry)

XX Peptide #9815 encoded by human foetal liver single exon probe.

XX Human; foetal liver; gene expression; single exon nucleic acid probe.

OS Homo sapiens.

PN WO200157277-A2.

XX 09-AUG-2001.

PF 30-JAN-2001; 2001WO-US000669.

XX 04-FEB-2000; 2000US-0180312P.

PR 26-MAY-2000; 2000US-0207456P.

PR 30-JUN-2000; 2000US-00608408.

PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human fetal liver.
XX
XX Claim 27; SEQ ID NO 34944; 639pp + Sequence Listing; English.
XX
CC The invention relates to a single exon nucleic acid probe for measuring
CC human gene expression in a sample derived from human foetal liver. The
CC single exon nucleic acid probes may be used for predicting, measuring and
CC displaying gene expression in samples derived from human fetal liver. The
CC present sequence is a peptide encoded by a single exon nucleic acid probe
CC of the invention. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 20 AA;
SQ
Query Match 100.0%; Score 36; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db |||||
3 HPPHG 7
RESULT 4
AAM36117
ID AAM36117 standard; protein; 20 AA.
XX
AC AAM36117;
XX
DT 17-OCT-2001 (first entry)
XX
DE Peptide #10154 encoded by probe for measuring placental gene expression.
XX
XX Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
OS Homo sapiens.
XX
PN WO200157272-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000663.
XX
XX 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488897/53.
XX
XX Human genome-derived single exon nucleic acid probes useful for analyzing
PT gene expression in human placenta.
XX

PS Claim 27; SEQ ID NO 36386; 654pp; English.
XX
CC The present invention relates to single exon nucleic acid probes (SENP:
CC see AA131315-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders
XX
SQ Sequence 20 AA;
Query Match 100.0%; Score 36; DB 4; Length 20;
Best Local Similarity 100.0%; Pred. No. 34;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db |||||
3 HPPHG 7
RESULT 5
ABB25810
ID ABB25810 standard; protein; 20 AA.
XX
AC ABB25810;
XX
DT 23-JAN-2002 (first entry)
XX
DE Protein #7809 encoded by probe for measuring heart cell gene expression.
XX
KW Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
PD 09-AUG-2001.
XX
PF 30-JAN-2001; 2001WO-US0000666.
XX
PR 04-FEB-2000; 2000US-0180312P.
PR 26-MAY-2000; 2000US-0207456P.
PR 30-JUN-2000; 2000US-00608408.
PR 03-AUG-2000; 2000US-00632366.
PR 21-SEP-2000; 2000US-0234687P.
PR 27-SEP-2000; 2000US-0236359P.
PR 04-OCT-2000; 2000GB-00024263.
XX
PA (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
XX
XX Claim 15; SEQ ID NO 27580; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting, the
CC diagnosing, grading, staging, monitoring and prognosing diseases of
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences

```

XX SQ Sequence 20 AA;
    Query Match      100.0%; Score 36; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 34;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 3 HPPHG 7

RESULT 6
AAM76009
ID AAM76009 standard; protein; 20 AA.
XX AC AAM76009;
XX DT 06-NOV-2001 (first entry)
XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 36315.
XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX OS Homo sapiens.
XX PN WO200157276-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000668.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
XX Human genome-derived single exon nucleic acid probes useful for analyzing
    gene expression in human bone marrow.
XX Example 4; SEQ ID NO 36315; 658pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
    probes which are derived from genomic sequences expressed in the human
    bone marrow. They can be used to measure gene expression in bone marrow
    samples, which may enable the improved diagnosis and treatment of cancers
    such as lymphoma, leukaemia and myeloma. The present sequence is a
    protein encoded by one of the probes of the invention

XX SQ Sequence 20 AA;
    Query Match      100.0%; Score 36; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 34;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 3 HPPHG 7

RESULT 7
AAM63195
ID AAM63195 standard; protein; 20 AA.

```

```

XX AC AAM63195;
XX DT 05-NOV-2001 (first entry)
XX DE Human brain expressed single exon probe encoded protein SEQ ID NO: 35300.
XX KW Human; brain expressed exon; gene expression analysis; probe; microarray;
KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
XX OS Homo sapiens.
XX PN WO200157275-A2.
XX PD 09-AUG-2001.
XX PF 30-JAN-2001; 2001WO-US000667.
XX PR 04-FEB-2000; 2000US-0180312P.
XX PR 26-MAY-2000; 2000US-0207456P.
XX PR 30-JUN-2000; 2000US-00608408.
XX PR 03-AUG-2000; 2000US-00632366.
XX PR 21-SEP-2000; 2000US-0234687P.
XX PR 27-SEP-2000; 2000US-0236359P.
XX PR 04-OCT-2000; 2000GB-00024263.
XX PA (MOLE-) MOLECULAR DYNAMICS INC.
XX PI Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483446/52.
XX Single exon nucleic acid probes for analyzing gene expression in human
    brains.
XX Example 4; SEQ ID NO 35300; 650pp + Sequence Listing; English.
XX The present invention provides a number of single exon nucleic acid
    probes which are derived from genomic sequences expressed in the human
    brain. They can be used to measure gene expression in brain cell samples,
    which may enable the diagnosis and improved treatment of nervous system
    diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
    epilepsy and cancers. The present sequence is a protein encoded by one of
    the probes of the invention

XX SQ Sequence 20 AA;
    Query Match      100.0%; Score 36; DB 4; Length 20;
    Best Local Similarity 100.0%; Pred. No. 34;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 3 HPPHG 7

RESULT 8
ABG57735
ID ABG57735 standard; peptide; 20 AA.
XX AC ABG57735;
XX DT 25-FEB-2003 (first entry)
XX DE Human liver peptide, SEQ ID No 36383.
XX KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
KW hypercholesterolaemia; coronary heart disease.
XX OS Homo sapiens.
XX PN WO200157273-A2.
XX

```

PD 09-AUG-2001.
 XX 30-JAN-2001; 2001WO-US000664.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2001-488898/53.
 DR Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 PT
 XX Claim 27; SEQ ID NO 36383; 558pp; English.
 XX The invention relates to a single exon nucleic acid probe (SENP) (I) for
 CC measuring human gene expression in a sample derived from human adult
 CC liver, comprising one of 13109 defined nucleotide sequences given in the
 CC specification (or complements/ fragments). The probe hybridises at high
 CC stringency to a nucleic acid molecule expressed in the human adult liver.
 CC (i) may be used for predicting, measuring and displaying gene expression
 CC in samples derived from human adult liver. The genes identified may be
 CC involved in genetic liver diseases such as cirrhosis,
 CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
 CC associated with coronary heart disease. ABG47348-ABG5930 represent human
 CC liver single exon encoded peptides of the invention. Note: The sequence
 CC information for this patent does not appear in the printed specification
 CC but was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 20 AA;
 SQ
 Query Match 100.0%; Score 36; DB 4; Length 20;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 DB |||||
 3 HPPHG 7
 RESULT 9
 ABG45414
 ID ABG45414 standard; peptide; 20 AA.
 XX
 XX ABG45414;
 AC
 XX
 DT 19-AUG-2002 (first entry)
 XX
 DE Human peptide encoded by genome-derived single exon probe SEQ ID 35079.
 XX
 KW Human; single exon probe; asthma; lung cancer; COPD; IID;
 KW chronic obstructive pulmonary disease; interstitial lung disease;
 KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
 KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
 KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
 KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagazer syndrome;
 KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
 KW primary ciliary dyskinesia; pulmonary hypertension;
 KW hyaline membrane disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200186003-A2.
 XX

PD 15-NOV-2001.
 XX 30-JAN-2001; 2001WO-US000665.
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 XX WPI; 2002-114183/15.
 DR Spatially-addressable set of single exon nucleic acid probes, used to
 PT measure gene expression in human lung samples.
 PT
 XX Claim 27; SEQ ID NO 35079; 634pp; English.
 XX The invention relates to a spatially-addressable set of single exon
 CC nucleic acid probes for measuring gene expression in a sample derived
 CC from human lung comprising single exon nucleic acid probes having one of
 CC 12614 nucleic acid sequences mentioned in the specification, or their
 CC complements or the 12387 open reading frames derived from the 12614
 CC probes. Also included are a microarray comprising the novel set of probes
 CC; the novel set of probes which hybridise at high stringency to a nucleic
 CC acid expressed in the human lung; measuring gene expression in a sample
 CC derived from human lung, comprising (a) contacting the array with a
 CC collection of detectably labeled nucleic acids derived from human lung
 CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karagazer syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 20 AA;
 Query Match 100.0%; Score 36; DB 5; Length 20;
 Best Local Similarity 100.0%; Pred. No. 34;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 DB |||||
 3 HPPHG 7

RESULT 10
AAU46033
ID AAU46033 standard; protein; 53 AA.
XX
XX
AC AAU46033;
XX
XX
DT 27-FEB-2002 (first entry)
XX
XX
DE Propionibacterium acnes immunogenic protein #6929.
XX
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.
XX
XX
OS Propionibacterium acnes.
XX
XX
PN WO200181581-A2.
XX
XX
PD 01-NOV-2001.
XX
XX
PF 20-APR-2001; 2001WO-US012865.
XX
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
XX
PA (CORI-) CORIXA CORP.
XX
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
XX
XX
DR WPI: 2001-616774/71.
DR N-PSDB; AAS95929.
XX
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
XX
PS Example 1; SEQ ID NO 7228; 1069pp; English.
XX
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 36; DB 4; Length 53;
Best Local Similarity 100.0%; Pred. NO. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 7 HPPHG 11

RESULT 11
ABM42552
ID ABM42552 standard; protein; 53 AA.
XX
XX
AC ABM42552;
XX
XX
DT 20-OCT-2003 (first entry)
XX
XX
DE Propionibacterium acnes predicted ORF-encoded polypeptide #7228.
XX
XX
KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX
XX
OS Propionibacterium acnes.
XX
XX
PN WO2003033515-A1.
XX
XX
PD 24-APR-2003.
XX
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
XX
PR 15-OCT-2001; 2001US-00978825.
XX
XX
PA (CORI-) CORIXA CORP.
XX
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglas J;
XX
XX
DR WPI: 2003-381789/36.
DR N-PSDB; ACF64458.
XX
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
XX
PS Example 1; SEQ ID NO 7228; 1481pp; English.
XX
XX
CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX
SQ Sequence 53 AA;

Query Match 100.0%; Score 36; DB 6; Length 53;
Best Local Similarity 100.0%; Pred. NO. 84;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

Db		7 HPPHG 11	Db		12 HPPHG 16
RESULT 12					
AAU54374					
ID	AAU54374 standard; protein; 61 AA.				
XX					
AC	AAU54374;				
XX					
DT	27-FEB-2002 (first entry)				
XX					
DE	Propionibacterium acnes immunogenic protein #15270.				
XX					
KW	SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;				
KW	uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;				
KW	inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;				
KW	dermatological; osteopathic; neuroprotectant.				
XX					
OS	Propionibacterium acnes.				
XX					
PN	WO200181581-A2.				
XX					
PD	01-NOV-2001.				
XX					
PF	20-APR-2001; 2001WO-US012865.				
XX					
PR	21-APR-2000; 2000US-0199047P.				
PR	02-JUN-2000; 2000US-0208841P.				
PR	07-JUL-2000; 2000US-0216747P.				
XX					
PA	(CORI-) CORIXA CORP.				
XX					
PI	Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;				
PI	L'Maisonneuve J, Zhang Y, Jen S, Carter D;				
XX					
DR	WPI; 2001-616774/71.				
DR	N-PSDB; AAS59564.				
XX					
PT	Propionibacterium acnes polypeptides and nucleic acids useful for				
PT	vaccinating against and diagnosing infections, especially useful for				
PT	treating acne vulgaris.				
XX					
PS	Example 1; SEQ ID NO 15569; 1069pp; English.				
XX					
CC	Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic				
CC	polypeptides. The proteins and their associated DNA sequences are used in				
CC	the treatment, prevention and diagnosis of medical conditions caused by				
CC	P. acnes. The disorders include SAPHO syndrome (synovitis, acne,				
CC	pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.				
CC	P. acnes is also involved in infections of bone, joints and the central				
CC	nervous system, however it is particularly involved in the inflammatory				
CC	lesions associated with acne vulgaris. A method for detecting the				
CC	presence or absence of P. acnes in a patient comprises contacting a				
CC	sample with a binding agent that binds to the proteins of the invention				
CC	and determining the amount of bound protein in the sample. The				
CC	polypeptides may be used as antigens in the production of antibodies				
CC	specific for P. acnes proteins. These antibodies can be used to				
CC	downregulate expression and activity of P. acnes polypeptides and				
CC	therefore treat P. acnes infections. The antibodies may also be used as				
CC	diagnostic agents for determining P. acnes presence, for example, by				
CC	enzyme linked immunosorbent assay (ELISA). Note: The sequence data for				
CC	this patent did not form part of the printed specification, but was				
CC	obtained in electronic format directly from WIPO at				
CC	ftp.wipo.int/pub/published_pct_sequences				
XX					
SQ	Sequence 61 AA;				
Query Match 100.0%; Score 36; DB 4; Length 61;					
Best Local Similarity 100.0%; Pred. No. 96;					
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
QY	1 HPPHG 5				
RESULT 13					
ABM50893					
ID	ABM50893 standard; protein; 61 AA.				
XX					
AC	ABM50893;				
XX					
DT	20-OCT-2003 (first entry)				
XX					
DE	Propionibacterium acnes predicted ORF-encoded polypeptide #15569.				
XX					
KW	Acne vulgaris; antisborrheic; dermatological; antibacterial;				
KW	immunostimulant; immune response; vaccine.				
XX					
OS	Propionibacterium acnes.				
XX					
PN	WO2003033515-A1.				
XX					
PD	24-APR-2003.				
XX					
PF	11-OCT-2002; 2002WO-US032727.				
XX					
PR	15-OCT-2001; 2001US-00978825.				
XX					
PA	(CORI-) CORIXA CORP.				
XX					
PI	Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;				
PI	Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;				
PI	Barth B, Vallieue-Douglas J;				
XX					
DR	WPI; 2003-381789/36.				
DR	N-PSDB; ACF64493.				
XX					
PT	New Propionibacterium acnes polypeptides and polynucleotides encoding the				
PT	polypeptide, useful for diagnosing, preventing or treating acne vulgaris,				
PT	or for stimulating an immune response specific for a P. acnes protein.				
XX					
PS	Example 1; SEQ ID NO 15569; 1481pp; English.				
XX					
CC	The invention relates to an isolated polynucleotide (ACF64435-ACF64733)				
CC	encoding a Propionibacterium acnes protein. The invention also relates to				
CC	polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to				
CC	immunogenic fragments of P. acnes polypeptides. The invention				
CC	additionally encompasses expression vectors and host cells comprising a				
CC	polynucleotide of the invention; antibodies against polypeptides of the				
CC	invention; fusion proteins comprising a polypeptide of the invention; a				
CC	method for stimulating an immune response specific for a P. acnes				
CC	polypeptide and an isolated T cell population comprising T cells prepared				
CC	via this method; a vaccine composition (comprising P. acnes polypeptides,				
CC	polynucleotides, antibodies, fusion proteins, T cell populations, or				
CC	antigen-presenting cells that express the polypeptide); a method and kit				
CC	for detecting or determining the presence or absence of P. acnes in a				
CC	patient; and a method for inhibiting the development of P. acnes in a				
CC	patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion				
CC	proteins, T cell populations or antigen-presenting cells that express the				
CC	polypeptides are useful for diagnosing, preventing or treating acne				
CC	vulgaris, or for stimulating an immune response specific for a P. acnes				
CC	protein. The polynucleotides can also be used as probes or primers for				
CC	nucleic acid hybridisation. The vaccine composition is useful for the				
CC	stimulation of an immune response against P. acnes, or for treating acne,				
CC	and the kit is useful for performing a diagnostic assay. The present				
CC	sequence represents a polypeptide predicted to be encoded by an ORF (open				
CC	reading frame) contained within the P. acnes polynucleotides of the				
CC	invention. Note: The sequence data for this patent did not form part of				
CC	the printed specification, but was obtained in electronic format directly				
CC	from WIPO at ftp.wipo.int/pub/published_pct_sequences				
XX					
SQ	Sequence 61 AA;				
Query Match 100.0%; Score 36; DB 6; Length 61;					

Best Local Similarity 100.0%; Pred. No. 96;			
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1 HPPHG 5		
Db	12 HPPHG 16		
RESULT 14			
AAU22059			
ID	AAU22059 standard; protein; 63 AA.		
AC	AAU22059;		
XX			
DT	18-DEC-2001 (first entry)		
XX			
DE	Human cardiovascular system antigen polypeptide SEQ ID No 833.		
XX			
KW	Cardiovascular system antigen; human; mouse; rabbit; goat; horse; cat;		
KW	chicken; sheep; immunosuppressive; antiarthritic; vasotropic; dog;		
KW	antihermetic; antiproliferative; cytostatic; cardiac; neuroprotective;		
KW	cerebroprotective; neotropic; antibacterial; virucide; fungicide; cancer;		
KW	ophthalmological; vulnary; gene therapy; autoimmune disease; neoplasm;		
KW	hyperproliferative disorder; breast; liver; cardiovascular disorder;		
KW	cerebrovascular disorder; nervous system disorder; bacterial infection;		
KW	fungal infection; viral infection; ocular disorder; endocrine disorder;		
KW	gastrointestinal disorder; renal disorder; respiratory disorder;		
KW	wound healing; skin aging; organ transplantation; tissue regeneration;		
KW	anti-infertility.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200155321-A2.		
XX			
PD	02-AUG-2001.		
XX			
PF	17-JAN-2001; 2001WO-US001340.		
XX			
PR	31-JAN-2000; 2000US-0179065P.		
PR	04-FEB-2000; 2000US-0180628P.		
PR	24-FEB-2000; 2000US-0184664P.		
PR	02-MAR-2000; 2000US-0186350P.		
PR	16-MAR-2000; 2000US-0189874P.		
PR	17-MAR-2000; 2000US-0190076P.		
PR	18-APR-2000; 2000US-0198123P.		
PR	19-MAY-2000; 2000US-0205515P.		
PR	07-JUN-2000; 2000US-0209467P.		
PR	28-JUN-2000; 2000US-0214886P.		
PR	30-JUN-2000; 2000US-0215135P.		
PR	07-JUL-2000; 2000US-0216647P.		
PR	07-JUL-2000; 2000US-0216880P.		
PR	11-JUL-2000; 2000US-0217487P.		
PR	11-JUL-2000; 2000US-0217496P.		
PR	14-JUL-2000; 2000US-0218230P.		
PR	26-JUL-2000; 2000US-0220963P.		
PR	26-JUL-2000; 2000US-0220964P.		
PR	14-AUG-2000; 2000US-0224518P.		
PR	14-AUG-2000; 2000US-0224519P.		
PR	14-AUG-2000; 2000US-0225213P.		
PR	14-AUG-2000; 2000US-0225214P.		
PR	14-AUG-2000; 2000US-0225266P.		
PR	14-AUG-2000; 2000US-0225267P.		
PR	14-AUG-2000; 2000US-0225268P.		
PR	14-AUG-2000; 2000US-0225270P.		
PR	14-AUG-2000; 2000US-0225447P.		
PR	14-AUG-2000; 2000US-0225757P.		
PR	14-AUG-2000; 2000US-0225758P.		
PR	14-AUG-2000; 2000US-0225759P.		
PR	18-AUG-2000; 2000US-0226279P.		
PR	22-AUG-2000; 2000US-0226681P.		
PR	22-AUG-2000; 2000US-0226686P.		
PR	22-AUG-2000; 2000US-0227182P.		
PR	23-AUG-2000; 2000US-0227009P.		

PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250361P.
 PR 01-DEC-2000; 2000US-0250391P.
 PR 05-DEC-2000; 2000US-0251030P.
 PR 05-DEC-2000; 2000US-0251988P.
 PR 05-DEC-2000; 2000US-0256719P.
 PR 06-DEC-2000; 2000US-0251479P.
 PR 08-DEC-2000; 2000US-0251856P.
 PR 08-DEC-2000; 2000US-0251868P.
 PR 08-DEC-2000; 2000US-0251869P.
 PR 08-DEC-2000; 2000US-0251989P.
 PR 08-DEC-2000; 2000US-0251990P.
 PR 11-DEC-2000; 2000US-0254097P.
 PR 05-JAN-2001; 2001US-0259678P.
 XX (HUMA-) HUMAN GENOME SCI INC.

XX PA
 XX PI Rosen CA, Barash SC, Ruben SM;
 XX DR
 XX DR WPI; 2001-451930/48.
 XX DR N-PSDB; AAS35333.

XX PT New cardiovascular system related polynucleotides and polypeptides,
 PT useful for diagnosing, treating and/or preventing disorders of the
 PT cardiovascular system.
 XX PS
 XX PS Claim 11; SEQ ID NO 833; 674pp; English.

XX CC Sequences AU21852-AU22466 represent the cardiovascular system antigen
 CC polypeptides of the invention. Cardiovascular system antigens and their
 CC associated polynucleotides are useful in the diagnosis, treatment and
 CC prevention of various types of disorders in e.g. humans, mice, rabbits,
 CC goats, horses, cats, dogs, chickens or sheep. A pathological condition
 CC can be determined by detecting the presence or absence of a mutation in a
 CC cardiovascular system antigen polynucleotide. The treatable disorders
 CC include autoimmune diseases such as rheumatoid arthritis,
 CC hyperproliferative disorders such as neoplasms of the breast or liver,
 CC cardiovascular disorders such as cardiac arrest, cerebrovascular
 CC disorders such as cerebral ischaemia, nervous system disorders such as
 CC Alzheimer's disease, infections caused by bacteria, viruses and fungi,
 CC ocular disorders such as corneal infection, endocrine disorders such as
 CC premature labour and infertility, gastrointestinal disorders such as
 CC Crohn's disease, renal disorders such as glomerulonephritis and
 CC respiratory disorders such as asthma and pleurisy. The polypeptides can
 CC also be used to aid wound healing, to prevent skin aging due to sunburn,
 CC to maintain organs before transplantation, to regenerate tissues and in
 CC chemotaxis. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

Query Match 100.0%; Score 36; DB 4; Length 63;
 Best Local Similarity 100.0%; Pred. No. 99;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 |||||
 Db 26 HPPHG 30

RESULT 15
 ADE46027
 ID ADE46027 standard; protein; 63 AA.
 XX AC ADE46027;
 XX DT 29-JAN-2004 (first entry)
 XX DE Human cardiovascular system related polypeptide #208.
 XX KW Human; cardiovascular system related polypeptide; cancer;
 KW proliferative disorder; foetal abnormality; developmental abnormality;
 KW haematopoietic disorder; AIDS; autoimmune disease; rheumatoid arthritis;
 KW inflammation; allergy; neurological disorder; Alzheimer's disease;
 KW Parkinson's disease; cognitive disorder; schizophrenia; asthma;
 KW skin disorder; psoriasis; sepsis; diabetes; atherosclerosis;
 KW cardiovascular disorder; angiotensin disorder; kidney disorder;
 KW gastrointestinal disorder; pregnancy-related disorder;
 KW endocrine disorder.
 XX OS Homo sapiens.
 XX PN US2003059908-A1.
 XX PD 27-MAR-2003.
 XX PF 07-MAR-2002; 2002US-00091504.
 XX PR 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216800P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 18-AUG-2000; 2000US-0225759P.
 PR 22-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0228668P.
 PR 23-AUG-2000; 2000US-0227182P.
 PR 30-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 06-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.

XX Human cardiovascular system associated polypeptide SeqID833.
 DE
 XX
 KW autoimmune disease; rheumatoid arthritis; hyperproliferative disorder;
 KW breast neoplasms; liver neoplasms; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; cerebral ischaemia;
 KW angiogenesis; nervous system disorder; Alzheimer's disease; infection;
 KW ocular disorder; corneal infection; wound healing;
 KW epithelial cell proliferation; skin aging; sunburn;
 KW organ transplantation; cell culture; tissue regeneration; chemotaxis;
 KW food additive; preservative; cardiovascular system associated antigen;
 KW nuclear factor kappaB; NFkappaB; promoter element; human.
 XX
 OS Homo sapiens.
 XX
 XX US2004005575-A1.
 XX
 XX 08-JAN-2004.
 XX
 XX 26-AUG-2002; 2002US-00227577.
 XX
 XX 31-JAN-2000; 2000US-0179065P.
 PR 04-FEB-2000; 2000US-0180628P.
 PR 24-FEB-2000; 2000US-0184664P.
 PR 02-MAR-2000; 2000US-0186350P.
 PR 16-MAR-2000; 2000US-0189874P.
 PR 17-MAR-2000; 2000US-0190076P.
 PR 18-APR-2000; 2000US-0198123P.
 PR 19-MAY-2000; 2000US-0205515P.
 PR 07-JUN-2000; 2000US-0209467P.
 PR 28-JUN-2000; 2000US-0214886P.
 PR 30-JUN-2000; 2000US-0215135P.
 PR 07-JUL-2000; 2000US-0216647P.
 PR 07-JUL-2000; 2000US-0216880P.
 PR 11-JUL-2000; 2000US-0217487P.
 PR 11-JUL-2000; 2000US-0217496P.
 PR 14-JUL-2000; 2000US-0218290P.
 PR 26-JUL-2000; 2000US-0220963P.
 PR 26-JUL-2000; 2000US-0220964P.
 PR 14-AUG-2000; 2000US-0224518P.
 PR 14-AUG-2000; 2000US-0224519P.
 PR 14-AUG-2000; 2000US-0225213P.
 PR 14-AUG-2000; 2000US-0225214P.
 PR 14-AUG-2000; 2000US-0225266P.
 PR 14-AUG-2000; 2000US-0225267P.
 PR 14-AUG-2000; 2000US-0225268P.
 PR 14-AUG-2000; 2000US-0225270P.
 PR 14-AUG-2000; 2000US-0225447P.
 PR 14-AUG-2000; 2000US-0225757P.
 PR 14-AUG-2000; 2000US-0225758P.
 PR 14-AUG-2000; 2000US-0225759P.
 PR 18-AUG-2000; 2000US-0226279P.
 PR 22-AUG-2000; 2000US-0226681P.
 PR 22-AUG-2000; 2000US-0226686P.
 PR 22-AUG-2000; 2000US-0227182P.
 PR 23-AUG-2000; 2000US-0227009P.
 PR 30-AUG-2000; 2000US-0228924P.
 PR 01-SEP-2000; 2000US-0229287P.
 PR 01-SEP-2000; 2000US-0229343P.
 PR 01-SEP-2000; 2000US-0229344P.
 PR 01-SEP-2000; 2000US-0229345P.
 PR 05-SEP-2000; 2000US-0229509P.
 PR 05-SEP-2000; 2000US-0229513P.
 PR 06-SEP-2000; 2000US-0230437P.
 PR 06-SEP-2000; 2000US-0230438P.
 PR 08-SEP-2000; 2000US-0231242P.
 PR 08-SEP-2000; 2000US-0231243P.
 PR 08-SEP-2000; 2000US-0231244P.
 PR 08-SEP-2000; 2000US-0231413P.
 PR 08-SEP-2000; 2000US-0231414P.
 PR 08-SEP-2000; 2000US-0232080P.
 PR 08-SEP-2000; 2000US-0232081P.
 PR 12-SEP-2000; 2000US-0231968P.
 PR 14-SEP-2000; 2000US-0232397P.
 PR 14-SEP-2000; 2000US-0232398P.
 PR 14-SEP-2000; 2000US-0232399P.
 PR 14-SEP-2000; 2000US-0232400P.
 PR 14-SEP-2000; 2000US-0232401P.
 PR 14-SEP-2000; 2000US-0233063P.
 PR 14-SEP-2000; 2000US-0233064P.
 PR 14-SEP-2000; 2000US-0233065P.
 PR 21-SEP-2000; 2000US-0234223P.
 PR 21-SEP-2000; 2000US-0234274P.
 PR 25-SEP-2000; 2000US-0234997P.
 PR 25-SEP-2000; 2000US-0234998P.
 PR 26-SEP-2000; 2000US-0235484P.
 PR 27-SEP-2000; 2000US-0235834P.
 PR 27-SEP-2000; 2000US-0235836P.
 PR 29-SEP-2000; 2000US-0236327P.
 PR 29-SEP-2000; 2000US-0236367P.
 PR 29-SEP-2000; 2000US-0236368P.
 PR 29-SEP-2000; 2000US-0236369P.
 PR 29-SEP-2000; 2000US-0236370P.
 PR 02-OCT-2000; 2000US-0236802P.
 PR 02-OCT-2000; 2000US-0237037P.
 PR 02-OCT-2000; 2000US-0237038P.
 PR 02-OCT-2000; 2000US-0237039P.
 PR 02-OCT-2000; 2000US-0237040P.
 PR 13-OCT-2000; 2000US-0239933P.
 PR 13-OCT-2000; 2000US-0239937P.
 PR 20-OCT-2000; 2000US-0240960P.
 PR 20-OCT-2000; 2000US-0241221P.
 PR 20-OCT-2000; 2000US-0241785P.
 PR 20-OCT-2000; 2000US-0241786P.
 PR 20-OCT-2000; 2000US-0241787P.
 PR 20-OCT-2000; 2000US-0241808P.
 PR 20-OCT-2000; 2000US-0241809P.
 PR 01-NOV-2000; 2000US-0244617P.
 PR 08-NOV-2000; 2000US-0246474P.
 PR 08-NOV-2000; 2000US-0246475P.
 PR 08-NOV-2000; 2000US-0246476P.
 PR 08-NOV-2000; 2000US-0246477P.
 PR 08-NOV-2000; 2000US-0246478P.
 PR 08-NOV-2000; 2000US-0246523P.
 PR 08-NOV-2000; 2000US-0246524P.
 PR 08-NOV-2000; 2000US-0246525P.
 PR 08-NOV-2000; 2000US-0246526P.
 PR 08-NOV-2000; 2000US-0246527P.
 PR 08-NOV-2000; 2000US-0246528P.
 PR 08-NOV-2000; 2000US-0246532P.
 PR 08-NOV-2000; 2000US-0246609P.
 PR 08-NOV-2000; 2000US-0246610P.
 PR 08-NOV-2000; 2000US-0246611P.
 PR 17-NOV-2000; 2000US-0249207P.
 PR 17-NOV-2000; 2000US-0249208P.
 PR 17-NOV-2000; 2000US-0249209P.
 PR 17-NOV-2000; 2000US-0249210P.
 PR 17-NOV-2000; 2000US-0249211P.
 PR 17-NOV-2000; 2000US-0249212P.
 PR 17-NOV-2000; 2000US-0249213P.
 PR 17-NOV-2000; 2000US-0249214P.
 PR 17-NOV-2000; 2000US-0249215P.
 PR 17-NOV-2000; 2000US-0249216P.
 PR 17-NOV-2000; 2000US-0249217P.
 PR 17-NOV-2000; 2000US-0249218P.
 PR 17-NOV-2000; 2000US-0249244P.
 PR 17-NOV-2000; 2000US-0249245P.
 PR 17-NOV-2000; 2000US-0249264P.
 PR 17-NOV-2000; 2000US-0249265P.
 PR 17-NOV-2000; 2000US-0249297P.
 PR 17-NOV-2000; 2000US-0249299P.
 PR 17-NOV-2000; 2000US-0249300P.
 PR 01-DEC-2000; 2000US-0250160P.
 PR 01-DEC-2000; 2000US-0250391P.

```
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764869.
PR 07-MAR-2002; 2002US-00091504.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2004-081713/08.
XX N-PSDB; ADJ06830.
XX
XX New cardiovascular system-related nucleic acid molecule, useful for
XX diagnosing, preventing or treating diseases of the cardiovascular system,
XX PT and in chromosome mapping, drug screening or in pharmacogenomics.
XX
XX Claim 11; SEQ ID NO 833; 262pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule encoding a
XX human cardiovascular system associated polypeptide (or antigens), or its
XX fragment. Also included recombinant vectors, recombinant host cells, an
XX isolated human cardiovascular system associated polypeptide (including
XX its fragment, allelic variant, species homologue or epitope), an isolated
XX antibody that binds specifically to a human cardiovascular system
XX associated polypeptide, diagnosing a pathological condition or
XX susceptibility to a pathological condition (comprising determining the
XX presence or absence of a mutation in human cardiovascular system
XX associated nucleic acid and diagnosing a condition based on the presence
XX or absence of the mutation), identifying a binding partner to human
XX cardiovascular system associated polypeptides, the gene corresponding to
XX the human cardiovascular system associated cDNA sequence and identifying
XX an activity in a biological assay comprising expressing the human
XX cardiovascular system associated cDNA in a cell, isolating the
XX supernatant, detecting an activity in a biological assay and identifying
XX the protein in the supernatant having the activity. The human
XX cardiovascular system associated nucleic acids and polypeptides are used
XX to prevent, treat or ameliorate a medical condition (for example in
XX humans, mice, rabbits, goats, horses, cats, dogs, chickens or sheep), for
XX example autoimmune diseases such as rheumatoid arthritis,
XX hyperproliferative disorders, for example neoplasms of the breast or
```

```
Query Match 100.0%; Score 36; DB 8; Length 63;
Best Local Similarity 100.0%; Pred. No. 99;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 HPPHG 5
Db 26 HPPHG 30
```

```
RESULT 17
AAU66176
ID AAU66176 standard; protein; 87 AA.
```

```
XX AAU66176;
```

```
XX 27-FEB-2002 (first entry)
```

```
XX Propionibacterium acnes immunogenic protein #27072.
```

```
XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
XX inflammatory lesion; acne vulgaris; immune response; vaccine;
XX dermatological; osteopathic; neuroprotectant.
```

```
XX OS Propionibacterium acnes.
XX WO200181581-A2.
XX 01-NOV-2001.
XX 20-APR-2001; 2001WO-US012865.
XX 21-APR-2000; 2000US-0199047P.
XX 02-JUN-2000; 2000US-0208841P.
XX 07-JUL-2000; 2000US-0216747P.
XX (CORI-) CORIXA CORP.
XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
XX L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX WPI; 2001-616774/71.
XX N-PSDB; AAS59717.
XX Propionibacterium acnes polypeptides and nucleic acids useful for
XX PT vaccinating against and diagnosing infections, especially useful for
XX PT treating acne vulgaris.
XX
XX Example 1; SEQ ID NO 27371; 1069pp; English.
XX
XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
XX polypeptides. The proteins and their associated DNA sequences are used in
XX the treatment, prevention and diagnosis of medical conditions caused by
XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
XX P. acnes is also involved in infections of bone, joints and the central
XX nervous system, however it is particularly involved in the inflammatory
XX lesions associated with acne vulgaris. A method for detecting the
XX presence or absence of P. acnes in a patient comprises contacting a
XX sample with a binding agent that binds to the proteins of the invention
XX and determining the amount of bound protein in the sample. The
XX polypeptides may be used as antigens in the production of antibodies
XX specific for P. acnes proteins. These antibodies can be used to
XX downregulate expression and activity of P. acnes polypeptides and
XX therefore treat P. acnes infections. The antibodies may also be used as
XX diagnostic agents for determining P. acnes presence, for example, by
XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
XX this patent did not form part of the printed specification, but was
XX obtained in electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 87 AA;
```

```
Query Match 100.0%; Score 36; DB 4; Length 87;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 1 HPPHG 5
Db 57 HPPHG 61
```

```
RESULT 18
ABM62695
ID ABM62695 standard; protein; 87 AA.
```

```
XX ABM62695;
```

```
XX 20-OCT-2003 (first entry)
```

```
XX Propionibacterium acnes predicted ORF-encoded polypeptide #27371.
```

```
XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
XX immunostimulant; immune response; vaccine.
XX Propionibacterium acnes.
```

XX W02003033515-A1.
 XX 24-APR-2003.
 XX 11-OCT-2002; 2002WO-US032727.
 XX 15-OCT-2001; 2001US-00978825.
 XX (CORI-) CORIXA CORP.
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 XX Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 XX Barth B, Vallieue-Douglas J;
 XX WPI; 2003-381789/36.
 XX N-PSDB; ACF64646.
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 XX polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 XX or for stimulating an immune response specific for a P. acnes protein.
 XX Example 1; SEQ ID NO 27371; 1481pp; English.
 XX The invention relates to an isolated polynucleotide (ACF64435-AC564733)
 XX encoding a Propionibacterium acnes protein. The invention also relates to
 XX polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 XX immunogenic fragments of P. acnes polypeptides. The invention
 XX additionally encompasses expression vectors and host cells comprising a
 XX polynucleotide of the invention; antibodies against polypeptides of the
 XX invention; fusion proteins comprising a polypeptide of the invention; a
 XX method for stimulating an immune response specific for a P. acnes
 XX polypeptide and an isolated T cell population comprising T cells prepared
 XX via this method; a vaccine composition (comprising P. acnes polypeptides,
 XX polynucleotides, antibodies, fusion proteins, T cell populations, or
 XX antigen-presenting cells that express the polypeptide); a method and kit
 XX for detecting or determining the presence or absence of P. acnes in a
 XX patient; and a method for inhibiting the development of P. acnes in a
 XX patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 XX proteins, T cell populations or antigen-presenting cells that express the
 XX polypeptides are useful for diagnosing, preventing or treating acne
 XX vulgaris, or for stimulating an immune response specific for a P. acnes
 XX protein. The polynucleotides can also be used as probes or primers for
 XX nucleic acid hybridisation. The vaccine composition is useful for the
 XX stimulation of an immune response against P. acnes, or for treating acne,
 XX and the kit is useful for performing a diagnostic assay. The present
 XX sequence represents a polypeptide predicted to be encoded by an ORF (open
 XX reading frame) contained within the P. acnes polynucleotides of the
 XX invention. Note: The sequence data for this patent did not form part of
 XX the printed specification, but was obtained in electronic format directly
 XX from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 87 AA;
 Query Match 100.0%; Score 36; DB 6; Length 87;
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db 57 HPPHG 61
 |||||
 |||||

RESULT 19
 AAU58530
 ID AAU58530 standard; protein; 92 AA.
 XX AC AAU58530;
 XX DT 27-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #19426.
 XX

XX KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 XX uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 XX inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant.
 XX OS Propionibacterium acnes.
 XX PN W0200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US012865.
 XX PR 21-APR-2000; 2000US-0199047P.
 XX PR 02-JUN-2000; 2000US-0208841P.
 XX PR 07-JUL-2000; 2000US-0216747P.
 XX (CORI-) CORIXA CORP.
 XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 XX PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 XX N-PSDB; AAS59592.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 XX vaccinating against and diagnosing infections, especially useful for
 XX treating acne vulgaris.
 XX Example 1; SEQ ID NO 19725; 1069pp; English.
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 XX polypeptides. The proteins and their associated DNA sequences are used in
 XX the treatment, prevention and diagnosis of medical conditions caused by
 XX P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 XX pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 XX P. acnes is also involved in infections of bone, joints and the central
 XX nervous system, however it is particularly involved in the inflammatory
 XX lesions associated with acne vulgaris. A method for detecting the
 XX presence or absence of P. acnes in a patient comprises contacting a
 XX sample with a binding agent that binds to the proteins of the invention
 XX and determining the amount of bound protein in the sample. The
 XX polypeptides may be used as antigens in the production of antibodies
 XX specific for P. acnes proteins. These antibodies can be used to
 XX downregulate expression and activity of P. acnes polypeptides and
 XX therefore treat P. acnes infections. The antibodies may also be used as
 XX diagnostic agents for determining P. acnes presence, for example, by
 XX enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 XX this patent did not form part of the printed specification, but was
 XX obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences

XX Sequence 92 AA;
 Query Match 100.0%; Score 36; DB 4; Length 92;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db 77 HPPHG 81
 |||||
 |||||

RESULT 20
 ABM55049
 ID ABM55049 standard; protein; 92 AA.
 XX AC ABM55049;
 XX DT 20-OCT-2003 (first entry)
 XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #19725.
 XX

KW Acne vulgaris; antisborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.
XX

OS Propionibacterium acnes.

PN WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieue-Douglass J;

XX WPI; 2003-381789/36.
XX N-PSDB; ACF64521.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX

XX Example 1; SEQ ID NO 19725; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC polynucleotide of the invention; fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide; a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 92 AA;

Query Match 100.0%; Score 36; DB 6; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

DB 77 HPPHG 81

RESULT 21

AAU63724

ID AAU63724 standard; protein; 94 AA.

XX AAU63724;

XX

DT 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #24620.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

XX 02-JUN-2000; 2000US-0208841P.

XX 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

XX N-PSDB; AAS59634.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.

XX Example 1; SEQ ID NO 24919; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX

XX Sequence 94 AA;

Query Match 100.0%; Score 36; DB 4; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

DB 64 HPPHG 68

RESULT 22

ABM60243

ID ABM60243 standard; protein; 94 AA.

XX ABM60243;

XX

DT 20-OCT-2003 (first entry)
 XX Propionibacterium acnes predicted ORF-encoded polypeptide #24919.
 DE Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 XX immunostimulant; immune response; vaccine.
 KW Propionibacterium acnes.
 XX WO2003033515-A1.
 XX 24-APR-2003.
 XX 11-OCT-2002; 2002WO-US032727.
 XX 15-OCT-2001; 2001US-00978825.
 XX (CORI-) CORIXA CORP.
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Valliève-Dougllass J;
 XX WPI; 2003-381789/36.
 DR N-PSDB; ACP64563.
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 PS Example 1; SEQ ID NO 24919; 1481pp; English.
 XX The invention relates to an isolated polynucleotide (ACF64435-ACP64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

ID AAU60014 standard; protein; 99 AA.
 XX AAU60014;
 AC 27-FEB-2002 (first entry)
 DT Propionibacterium acnes immunogenic protein #20910.
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX Propionibacterium acnes.
 OS WO200181581-A2.
 XX 01-NOV-2001.
 XX 20-APR-2001; 2001WO-US012865.
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX (CORI-) CORIXA CORP.
 PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX WPI; 2001-616774/71.
 DR N-PSDB; AAS59607.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 PS Example 1; SEQ ID NO 21209; 1069pp; English.
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: the sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

ID ABB15846 standard; protein; 99 AA.
AC ABB15846;
XX
XX
DT 23-JAN-2002 (first entry)
XX
DE Human nervous system related polypeptide SEQ ID NO 4503.
XX
XX Human; nootropic; neuroprotective; cytostatic; dermatological; virucide;
KW immunosuppressive; antiinflammatory; anti-HIV; antibacterial; vulnerary;
KW antiparkinsonian; antiskinking; antianaemic; antiarthritic; cancer;
KW antineumatic; hepatotropic; cerebroprotective; antiinflammatory;
KW antiallergic; antidiabetic; antiulcer; anticonvulsant; antifungal;
KW antiparasitic; cardiant; immune disorder; cardiovascular disorder;
KW neurological disease; infection; nephrotropic; gene therapy; vaccine.
XX
OS Homo sapiens.
XX
XX WO200159063-A2.
XX
XX 16-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001334.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.

KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
KW antipsoriatic; dermatological; cardiant; vasotropic; vulnerary;
KW angiogenesis; gene therapy; rabbit.
XX
OS Oryctolagus cuniculus.
XX
PN WO2003077872-A2.
XX
XX 25-SEP-2003.
XX
XX 17-MAR-2003; 2003WO-US008060.
PF
XX 15-MAR-2002; 2002US-0364047P.
PR
XX (ATTE-) ATTENUON LLC.
PA
XX McCrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.
DR
XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
PT cancer.
XX
XX Claim 6; SEQ ID NO 24; 117bp; English.
PS
XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents: The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding, HK or its D5
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumours or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a rabbit HPRG
CC protein His-Pro (H/P) rich domain.
XX
SQ Sequence 101 AA;
Query Match 100.0%; Score 36; DB 8; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.6e+02; Mismatches 0; Gaps 0;
Matches 5; Conservative 0; Indels 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 33 HPPHG 37
|||
|||
RESULT 27
AAO13856
ID AAO13856 standard; protein; 112 AA.
XX
XX AAO13856;
AC
XX
XX 06-NOV-2001 (first entry)
DT
XX

DE Human polypeptide SEQ ID NO 27748.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
XX 07-SEP-2001.
PD
XX 26-FEB-2001; 2001WO-US004927.
PF
XX 28-FEB-2000; 2000US-00515126.
PR
XX 18-MAY-2000; 2000US-00577409.
PR
XX (HYSE-) HYSEQ INC.
PA
XX Tang YT, Liu C, Drmanac RT;
PI
XX WPI; 2001-514838/56.
DR
XX N-FSDS; AAI93787.
DR
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX Claim 20; SEQ ID NO 27748; 1399pp + Sequence Listing; English.
PS
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 112 AA;
Query Match 100.0%; Score 36; DB 4; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Mismatches 0; Gaps 0;
Matches 5; Conservative 0; Indels 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 96 HPPHG 100
|||
|||
RESULT 28
ABG70306
ID ABG70306 standard; protein; 117 AA.
XX
XX AC ABG70306;
XX
XX 21-OCT-2002 (first entry)
DT
XX Human MDDT protein Incyte ID No: LI:180252.16.orf2:2001JAN12.
DE
XX Human; molecule for disease detection and treatment; MDDT; cancer;
KW cell proliferative disorder; arteriosclerosis; cirrhosis; asthma;
KW autoimmune disorder; inflammatory disorder; Crohn's disease;
KW multiple sclerosis; cytostatic; antiarteriosclerotic; antiinflammatory;
KW hepatotropic; immunosuppressive; antiasthmatic.
XX
XX Homo sapiens.
OS
XX

PN WO200255738-A2.
 XX 18-JUL-2002.
 PD
 XX 09-JAN-2002; 2002WO-US001008.
 XX 12-JAN-2001; 2001US-0261622P.
 PR 16-JAN-2001; 2001US-0261865P.
 PR 17-JAN-2001; 2001US-0262208P.
 PR 17-JAN-2001; 2001US-0262209P.
 PR 17-JAN-2001; 2001US-0262326P.
 PR 19-JAN-2001; 2001US-0263063P.
 PR 19-JAN-2001; 2001US-0263065P.
 PR 19-JAN-2001; 2001US-0263329P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Panzer SR, Lincoln SE, Altus CM, Dufour GE, Hillman JL, Jones AL;
 PI Dam TC, Liu TF, Harris B, Flores V, Daffo A, Marwaha R, Chen AJ;
 PI Chang SC, Gerstin EH, Peralta CH, David MH, Lewis SA;
 XX WPI; 2002-590679/63.
 DR N-PSDB; ABS51779.
 XX
 XX New disease detection and treatment molecule (MDDT) polynucleotides and
 PT polypeptides, useful in diagnosing, studying, preventing or treating
 PT diseases associated with MDDT expression, e.g. autoimmune or inflammatory
 PT disorders.
 XX
 PS Claim 27; Page 107; 129pp; English.
 CC The present invention relates to the isolation of novel human molecules
 CC for disease detection and treatment (MDDT), and the polynucleotide
 CC sequences (mdmt) encoding them. The MDDT polypeptides may be used to
 CC screen for molecules that bind to, or are bound by the encoded
 CC polypeptides, and to develop a transcript image of a tissue or cell type.
 CC Probes comprising at least 20 nucleotides of the mdmt polynucleotide may
 CC be used to assess the toxicity of a test compound. The MDDT polypeptides
 CC and mdmt polynucleotides are useful in the diagnosis, study, prevention
 CC and treatment of diseases associated with the expression of molecules for
 CC disease detection and treatment. Such disorders include cell
 CC proliferative disorders (e.g. arteriosclerosis, cirrhosis, or cancers),
 CC and autoimmune/inflammatory disorders (e.g. asthma, Crohn's disease, or
 CC multiple sclerosis). The mdmt polynucleotides may also be used as
 CC molecule markers, in microarrays, and in somatic or germline gene
 CC therapy. ABG70306-ABG70341 represent the MDDT proteins of the invention
 XX
 XX Sequence 117 AA;
 SQ
 Query Match 100.0%; Score 36; DB 5; Length 117;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 |||||
 Db 17 HPPHG 21
 RESULT 29
 AAU00910
 ID AAU00910 standard; protein; 118 AA.
 XX
 AC AAU00910;
 DT
 DT 04-JUL-2001 (first entry)
 DE Human cancer related protein 25#1.
 XX
 XX Human; cancer related protein; food additive; preservative; immunogen;
 KW antibody; bone cancer; adrenal cancer; bone marrow cancer; breast cancer;
 KW gastrointestinal cancer; liver cancer; lung cancer; urogenital cancer;
 KW immune disorder; Addison's disease; allergy; diabetes mellitus;
 KW autoimmune haemolytic anaemia; autoimmune thyroiditis; Crohn's disease;

KW multiple sclerosis; rheumatoid arthritis; ulcerative colitis;
 KW acquired immunodeficiency syndrome; AIDS; cardiovascular disorder;
 KW myocardial ischaemia; wound healing; neurological disorder;
 KW Parkinson's disease; Alzheimer's disease; cerebral anoxia; epilepsy;
 KW viral infection; bacterial infection; fungal infection;
 KW parasitic infection; agonist; antagonist.
 XX
 OS Homo sapiens.
 XX
 XX WO200118014-A1.
 PN
 XX 15-MAR-2001.
 PD
 XX 30-AUG-2000; 2000WO-US023794.
 PF
 XX 03-SEP-1999; 99US-0152296P.
 PR 06-OCT-1999; 99US-0158003P.
 PR
 XX (HUMA-) HUMAN GENOME SCI INC.
 PA
 XX Roschke V;
 PI
 XX WPI; 2001-235186/24.
 DR N-PSDB; AAS00851.
 DR
 XX Twenty nine nucleic acid molecules encoding human cancer associated
 PT proteins, useful in the prevention, treatment and diagnosis of cancer,
 PT immune disorders, cardiovascular disorders and neurological diseases.
 PT
 XX Disclosure; Page 422-423; 427pp; English.
 PS
 XX The sequence represents a novel Human cancer related protein. The
 CC polynucleotides and polypeptides are useful for preventing, treating or
 CC ameliorating a medical condition in e.g. humans, mice, rabbits, goats,
 CC horses, cats, dogs, chickens or sheep. The polypeptides can also be used
 CC as a food additive or preservative to increase or decrease storage
 CC capabilities. The polynucleotide are useful for chromosome
 CC identification. The nucleic acids, protein, antibodies, agonists and
 CC antagonists are useful in the diagnosis, treatment and prevention of
 CC cancer (e.g. cancers of the adrenal gland, bone, bone marrow, breast,
 CC gastrointestinal tract, liver, lung, or urogenital), immune disorders
 CC (e.g. Addison's disease, allergies, autoimmune haemolytic anaemia,
 CC autoimmune thyroiditis, diabetes mellitus, Crohn's disease, multiple
 CC sclerosis, rheumatoid arthritis and ulcerative colitis, acquired
 CC immunodeficiency syndrome, AIDS), cardiovascular disorders such as
 CC myocardial ischaemia, wound healing, neurological diseases (e.g.
 CC Parkinson's disease, Alzheimer's disease, cerebral anoxia and epilepsy)
 CC and infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Numerous examples of each type of disorder are given in the
 CC specification
 XX
 SQ Sequence 118 AA;
 Query Match 100.0%; Score 36; DB 4; Length 118;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 |||||
 Db 96 HPPHG 100
 RESULT 30
 AAU40328
 ID AAU40328 standard; protein; 202 AA.
 XX
 AC AAU40328;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #1224.
 XX
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertonosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 OS
 OS Propionibacterium acnes.
 XX
 XX WO200181581-A2.
 PN
 PD
 PD 01-NOV-2001.
 XX
 XX 20-APR-2001; 2001WO-US012865.
 PF
 PF 21-APR-2000; 2000US-0199047P.
 PR
 PR 02-JUN-2000; 2000US-0208841P.
 PR
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'Maisonneuve J, Zhang Y, Jen S, Carter D;
 PI N-PSDB; AAS59511.
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59511.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 PT
 PS Example 1; SEQ ID NO 1523; 1069pp; English.
 XX
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 202 AA;
 Query Match 100.0%; Score 36; DB 4; Length 202;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db |||||
 93 HPPHG 97
 RESULT 31
 ABM36847
 ID ABM36847 standard; protein; 202 AA.
 XX
 AC ABM36847;
 XX
 XX 20-OCT-2003 (first entry)
 DT
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #1523.
 XX
 XX Acne vulgaris; antisieborrhoic; dermatological; antibacterial;
 KW

KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 XX WO2003033515-A1.
 PN
 PD 24-APR-2003.
 XX
 XX 11-OCT-2002; 2002WO-US032727.
 PF
 PF 15-OCT-2001; 2001US-00978825.
 PR
 PR (CORI-) CORIXA CORP.
 PA
 XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglas J;
 XX WPI; 2003-381789/36.
 DR N-PSDB; ACF64440.
 DR
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 1523; 1481pp; English.
 XX
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 202 AA;
 Query Match 100.0%; Score 36; DB 6; Length 202;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db |||||
 93 HPPHG 97
 RESULT 32
 AAY29941
 ID AAY29941 standard; protein; 231 AA.
 XX
 AC AAY29941;
 XX
 DT 22-NOV-1999 (first entry)

XX DE Zea mays pathogenesis-related class I PR-1#70 protein.
XX KW Zea mays; maize; pathogenesis-related class I; PR-1; promoter;
KW regulation; expression; disease resistance; genetic manipulation;
KW tobacco mosaic virus; cucumber mosaic virus; ringspot virus;
KW necrosis virus; maize dwarf virus; viroid; bacterial; insect; nematode;
KW fungal.
XX OS Zea mays.
XX PN WO9943819-A1.
XX PD 02-SEP-1999.
XX PF 11-FEB-1999; 99WO-US003011.
XX PR 26-FEB-1998; 98US-0076100P.
XX PR 27-MAR-1998; 98US-0079648P.
XX PA (PION-) PIONEER HI-BRED INT INC.
XX PI Crane VC;
XX PI related PR-1 proteins. The promoters are useful for expressing
XX heterologous genes (including genes for disease resistance) in plants,
XX especially dicots, or monocots i.e. maize. The promoters are useful for
XX the genetic manipulation of plants to exhibit specific phenotypes,
XX particularly enhanced resistance to pathogen-caused disease. Pathogens
XX include viruses such as tobacco or cucumber mosaic virus, ringspot virus,
XX necrosis virus, and maize dwarf virus, and viroids, bacteria, insects,
XX nematodes and fungi. The present sequence represents a maize PR-1 protein
XX given in the present invention
XX SQ Sequence 231 AA;
XX
XX Query Match 100.0%; Score 36; DB 2; Length 231;
XX Best Local Similarity 100.0%; Pred. No. 3.4e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 HPPHG 5
DB 174 HPPHG 178
RESULT 33
ADA48350
ID ADA48350 standard; protein; 241 AA.
XX
XX ADA48350;
XX
XX 20-NOV-2003 (first entry)
XX
XX Rice protein conferring disease resistance in plants.
XX disease resistance; pathogen tolerance; plant pathogen; plant; rice.
XX Oryza sativa.
XX
XX WO2003000906-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-IB002453.
XX PF

XX 22-JUN-2001; 2001US-0300112P.
XX PR 26-SEP-2001; 2001US-0352277P.
XX PR 22-MAR-2002; 2002US-0366535P.
XX
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Glazebrook J, Briggs S, Cooper B, Goff SA, Moughamer T;
PI Katagiri F, Kreps J, Provart N, Ricke D, Zhu T;
XX
XX WPI: 2003-184052/18.
DR N-PSDB; ADA48349.
XX
XX New polynucleotide comprising a plant nucleotide sequence having an open
PT reading frame that encodes a polypeptide associated with disease
PT resistance, useful for conferring resistance or tolerance to a plant
PT pathogen.
XX
XX Claim 10; SEQ ID NO 420; 299pp; English.
XX
XX The invention relates to a novel isolated polynucleotide comprising a
CC plant nucleotide sequence having an open reading frame that encodes a
CC polypeptide associated with disease resistance or its fragment having
CC substantially the same activity as the full-length polypeptide. The
CC polynucleotide of the invention is useful for conferring resistance or
CC tolerance to a plant pathogen. The present sequence represents a protein
CC conferring disease resistance used in the invention.
XX
XX SQ Sequence 241 AA;
XX
XX Query Match 100.0%; Score 36; DB 6; Length 241;
XX Best Local Similarity 100.0%; Pred. No. 3.5e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
QY 1 HPPHG 5
DB 178 HPPHG 182
RESULT 34
AAU87400
ID AAU87400 standard; protein; 262 AA.
XX
XX AAU87400;
XX
XX 05-JUN-2002 (first entry)
XX
XX Novel central nervous system protein #310.
XX
XX Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischaemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy.
XX
XX Homo sapiens.
OS
XX
XX WO200155318-A2.
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US001332.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 24-FEB-2000; 2000US-0184664P.
XX PR 02-MAR-2000; 2000US-0186350P.
XX PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0250300P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 11-DEC-2000; 2000US-0251990P.
PR 05-JAN-2001; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
PA Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-581633/65.
DR N-PSDB; ABK43730.
XX
PT New isolated nucleic acid encoding a protein for diagnosing, preventing, or
PT treating or ameliorating medical conditions and used as food additives or
XX preservatives.
XX
PS Claim 9; SEQ ID NO 918; 837pp; English.
XX The invention describes an isolated nucleic acid molecule (I) encoding a

CC novel central nervous system protein. (I) and polypeptides (III) encoded
 CC by (I), are used to treat a medical conditions and in diagnosis of a
 CC pathological condition. Disorders which are diagnosed or treated include
 CC autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC anglogenesis, nervous system disorders e.g. Alzheimer's disease and
 CC amyotrophic lateral sclerosis, infections caused by bacteria, viruses
 CC e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders
 CC e.g. corneal infection, gastrointestinal disorders e.g. dysphagia,
 CC adenocarcinomas and irritable bowel syndrome, reproductive system
 CC disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes
 CC and pituitary dwarfism, cancers and disorders at the cellular level e.g.
 CC leukaemia, disorders involving neovascularisation e.g. malignancies,
 CC respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g.
 CC acute kidney failure and blood related disorders e.g. myocardial
 CC infarction. The polypeptides can also be used to aid wound healing and
 CC epithelial cell proliferation, to prevent skin aging due to sunburn, to
 CC maintain organs before transplantation, for supporting cell culture of
 CC primary tissues, to regenerate tissues and in chemotaxis. The
 CC polypeptides can also be used as a food additive or preservative to
 CC increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 100.0%; Score 36; DB 4; Length 262;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;

QY 1 HPPHG 5
 Db 17 HPPHG 21

RESULT 35

ABO61857
 ID ABO61857 standard; protein; 262 AA.

AC ABO61857;

DT 29-JUL-2004 (first entry)

DE Klebsiella pneumoniae polypeptide seqid 8374.

XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; vaccine.

OS Klebsiella pneumoniae.

PN US6610836-B1.

XX 26-AUG-2003.

PF 27-JAN-2000; 2000US-00489039.

XX 29-JAN-1999; 99US-0117747P.

PR (GENO-) GENOME THERAPEUTICS CORP.

PA Breton GL, Osborne M;

XX WPI: 2003-895346/82.

DR N-PSDB; ACH95408.

XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.

PS Disclosure; SEQ ID NO 8374; 932pp; English.

XX The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a

CC Klebsiella pneumoniae polypeptide of the invention
 XX Sequence 262 AA;

Query Match 100.0%; Score 36; DB 7; Length 262;
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 Db 18 HPPHG 22

RESULT 36

AD154715
 ID AD154715 standard; protein; 262 AA.

XX AC AD154715;

DT 22-APR-2004 (first entry)

XX Novel human protein seq id 918.

XX neuroprotective; nootropic; antiparkinsonian; anticonvulsant;
 KW antidiabetic; antirheumatic; antiarthritic; dermatological;
 KW antiinflammatory; immunosuppressive; antithyroid; antianemic; vasotropic;
 KW anti-HIV; hepatotropic; virucide; antibacterial; fungicide;
 KW antiparasitic; muscular; gynaecological; gastrointestinal; respiratory;
 KW cardiovascular; litholytic; cytostatic; gene therapy; neural disorder;
 KW Alzheimer's disease; Parkinson's disease; Huntington's chorea;
 KW amyotrophic lateral sclerosis; multiple sclerosis;
 KW immune system disorder; diabetes; rheumatoid arthritis;
 KW systemic lupus erythematosus; autoimmune thyroiditis;
 KW inflammatory bowel disease; ischaemia-reperfusion injury;
 KW HIV infection; hepatitis infection; bacterial infection;
 KW fungal infection; parasitic infection; muscular disorder;
 KW reproductive disorder; gastrointestinal disorder; pulmonary disorder;
 KW cardiovascular disorder; atherosclerosis; arrhythmia; myocarditis;
 KW renal disorder; acute glomerulonephritis; pyelonephritis;
 KW renal lithiasis; proliferative disorder; cancerous diseases; human.

OS Homo sapiens.

PN US2004018969-A1.

PD 29-JAN-2004.

XX 17-JAN-2001; 2001US-00764875.

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0216647P.

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR	14-AUG-2000;	2000US-0225266P.	PR	08-NOV-2000;	2000US-0246525P.
PR	14-AUG-2000;	2000US-0225267P.	PR	08-NOV-2000;	2000US-0246526P.
PR	14-AUG-2000;	2000US-0225268P.	PR	08-NOV-2000;	2000US-0246527P.
PR	14-AUG-2000;	2000US-0225270P.	PR	08-NOV-2000;	2000US-0246528P.
PR	14-AUG-2000;	2000US-0225447P.	PR	08-NOV-2000;	2000US-0246532P.
PR	14-AUG-2000;	2000US-0225757P.	PR	08-NOV-2000;	2000US-0246609P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246610P.
PR	14-AUG-2000;	2000US-0225758P.	PR	08-NOV-2000;	2000US-0246611P.
PR	18-AUG-2000;	2000US-0226279P.	PR	08-NOV-2000;	2000US-0246613P.
PR	22-AUG-2000;	2000US-0226681P.	PR	17-NOV-2000;	2000US-0249207P.
PR	22-AUG-2000;	2000US-0226868P.	PR	17-NOV-2000;	2000US-0249208P.
PR	23-AUG-2000;	2000US-0227182P.	PR	17-NOV-2000;	2000US-0249209P.
PR	23-AUG-2000;	2000US-0227009P.	PR	17-NOV-2000;	2000US-0249210P.
PR	30-AUG-2000;	2000US-0228924P.	PR	17-NOV-2000;	2000US-0249211P.
PR	01-SEP-2000;	2000US-0229287P.	PR	17-NOV-2000;	2000US-0249212P.
PR	01-SEP-2000;	2000US-0229343P.	PR	17-NOV-2000;	2000US-0249213P.
PR	01-SEP-2000;	2000US-0229344P.	PR	17-NOV-2000;	2000US-0249214P.
PR	01-SEP-2000;	2000US-0229345P.	PR	17-NOV-2000;	2000US-0249215P.
PR	05-SEP-2000;	2000US-0229509P.	PR	17-NOV-2000;	2000US-0249216P.
PR	05-SEP-2000;	2000US-0229513P.	PR	17-NOV-2000;	2000US-0249217P.
PR	06-SEP-2000;	2000US-0230437P.	PR	17-NOV-2000;	2000US-0249218P.
PR	08-SEP-2000;	2000US-0230438P.	PR	17-NOV-2000;	2000US-0249244P.
PR	08-SEP-2000;	2000US-0231242P.	PR	17-NOV-2000;	2000US-0249245P.
PR	08-SEP-2000;	2000US-0231243P.	PR	17-NOV-2000;	2000US-0249264P.
PR	08-SEP-2000;	2000US-0231244P.	PR	17-NOV-2000;	2000US-0249265P.
PR	08-SEP-2000;	2000US-0231413P.	PR	17-NOV-2000;	2000US-0249297P.
PR	08-SEP-2000;	2000US-0231414P.	PR	17-NOV-2000;	2000US-0249299P.
PR	08-SEP-2000;	2000US-0232080P.	PR	17-NOV-2000;	2000US-0249300P.
PR	08-SEP-2000;	2000US-0232081P.	PR	01-DEC-2000;	2000US-0250160P.
PR	12-SEP-2000;	2000US-0231968P.	PR	01-DEC-2000;	2000US-0250391P.
PR	14-SEP-2000;	2000US-0232397P.	PR	05-DEC-2000;	2000US-0251030P.
PR	14-SEP-2000;	2000US-0232398P.	PR	05-DEC-2000;	2000US-0251988P.
PR	14-SEP-2000;	2000US-0232399P.	PR	05-DEC-2000;	2000US-0256719P.
PR	14-SEP-2000;	2000US-0232400P.	PR	08-DEC-2000;	2000US-0251479P.
PR	14-SEP-2000;	2000US-0232401P.	PR	08-DEC-2000;	2000US-0251856P.
PR	14-SEP-2000;	2000US-0233063P.	PR	08-DEC-2000;	2000US-0251868P.
PR	14-SEP-2000;	2000US-0233064P.	PR	08-DEC-2000;	2000US-0251869P.
PR	14-SEP-2000;	2000US-0233065P.	PR	08-DEC-2000;	2000US-0251989P.
PR	21-SEP-2000;	2000US-0234223P.	PR	08-DEC-2000;	2000US-0251990P.
PR	21-SEP-2000;	2000US-0234274P.	PR	11-DEC-2000;	2000US-0254097P.
PR	25-SEP-2000;	2000US-0234997P.	PR	05-JAN-2001;	2001US-0259678P.
PR	25-SEP-2000;	2000US-0234998P.	XX		
PR	26-SEP-2000;	2000US-0235484P.	PA	(ROSE/) ROSEN C A.	
PR	27-SEP-2000;	2000US-0235834P.	PA	(RUBE/) RUBEN S M.	
PR	27-SEP-2000;	2000US-0235836P.	PA	(BARA/) BARASH S C.	
PR	29-SEP-2000;	2000US-0236327P.	XX		
PR	29-SEP-2000;	2000US-0236367P.	PI	Rosen CA, Ruben SM, Barash SC;	
PR	29-SEP-2000;	2000US-0236368P.	XX		
PR	29-SEP-2000;	2000US-0236369P.	DR	WPI; 2004-122079/12.	
PR	29-SEP-2000;	2000US-0236370P.	DR	N-PSDB; ADI54117.	
PR	02-OCT-2000;	2000US-0236802P.	XX		
PR	02-OCT-2000;	2000US-0237037P.	PT	New polypeptides and nucleic acid molecules, useful for detecting,	
PR	02-OCT-2000;	2000US-0237038P.	PT	preventing, diagnosing, prognosticating, treating or ameliorating medical	
PR	02-OCT-2000;	2000US-0237039P.	PT	conditions e.g. neural disorders, reproductive disorders or infectious	
PR	02-OCT-2000;	2000US-0237040P.	PT	diseases.	
PR	13-OCT-2000;	2000US-0239935P.	XX		
PR	13-OCT-2000;	2000US-0239935P.	XX	Claim 11; SEQ ID NO 918; 413pp; English.	
PR	20-OCT-2000;	2000US-0240960P.	XX		
PR	20-OCT-2000;	2000US-0241221P.	CC	The invention describes an isolated polypeptide comprising an amino acid	
PR	20-OCT-2000;	2000US-0241785P.	CC	sequence at least 90% identical to: a polypeptide fragment, domain,	
PR	20-OCT-2000;	2000US-0241786P.	CC	epitope, or full-length protein of any one of 607 amino acid sequences	
PR	20-OCT-2000;	2000US-0241787P.	CC	(I) described in the specification; a polypeptide fragment of (I), or the	
PR	20-OCT-2000;	2000US-0241808P.	CC	encoded sequence contained in (II), having biological activity; or a	
PR	20-OCT-2000;	2000US-0241809P.	CC	variant, allelic variant, or a species homologue of (I). The polypeptides	
PR	20-OCT-2000;	2000US-0241809P.	CC	and nucleic acid molecules are useful for detecting, preventing,	
PR	01-NOV-2000;	2000US-0241826P.	CC	diagnosing, prognosticating, treating or ameliorating medical conditions	
PR	08-NOV-2000;	2000US-0246474P.	CC	such as neural disorders, e.g. Alzheimer's disease, Parkinson's disease,	
PR	08-NOV-2000;	2000US-0246475P.	CC	Huntington's chorea, amyotrophic lateral sclerosis or multiple sclerosis,	
PR	08-NOV-2000;	2000US-0246476P.	CC	immune system disorders, e.g. diabetes, rheumatoid arthritis, systemic	
PR	08-NOV-2000;	2000US-0246478P.			
PR	08-NOV-2000;	2000US-0246523P.			
PR	08-NOV-2000;	2000US-0246524P.			
Query Match			100.0%; Score 36; DB 8; Length 262;		
Best Local Similarity			100.0%; Pred. No. 3.8e+02;		
Matches			5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

QY 1 HPPHG 5
 Db 17 HPPHG 21
 RESULT 37
 ABB66594
 ID ABB66594 standard; protein; 311 AA.
 XX ABB66594;
 AC ABB66594;
 DT 26-MAR-2002 (first entry)
 DE Drosophila melanogaster polypeptide SEQ ID NO 26574.
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX Drosophila melanogaster.
 OS WO200171042-A2.
 PN 27-SEP-2001.
 XX 23-MAR-2001; 2001WO-US009231.
 XX 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX (PEKE) PE CORP NY.
 XX Venter JC, Adams M, Li PWD, Myers EW;
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL10697.
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX Disclosure; SEQ ID NO 26574; 21pp + Sequence Listing; English.
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO.int/pub/published_pct_sequences
 XX Sequence 311 AA;
 SQ
 Query Match 100.0%; Score 36; DB 4; Length 311;
 Best Local Similarity 100.0%; Pred. No. 4.5e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db 102 HPPHG 106
 RESULT 38
 AAU51225
 ID AAU51225 standard; protein; 324 AA.
 XX AAU51225;
 AC AAU51225;
 XX 27-FEB-2002 (first entry)
 DT
 XX

DE Propionibacterium acnes immunogenic protein #12121.
 XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX Propionibacterium acnes.
 OS WO200181581-A2.
 PN 01-NOV-2001.
 XX 20-APR-2001; 2001WO-US012865.
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX (CORI-) CORIXA CORP.
 PA Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 PI WPI; 2001-616774/71.
 DR N-PSDB; AAS59551.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX Example 1; SEQ ID NO 12420; 1069pp; English.
 PS Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 XX polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX Sequence 324 AA;
 SQ
 Query Match 100.0%; Score 36; DB 4; Length 324;
 Best Local Similarity 100.0%; Pred. No. 4.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 Db 254 HPPHG 258
 RESULT 39
 ABM47744
 ID ABM47744 standard; protein; 334 AA.
 XX ABM47744;
 AC ABM47744;
 XX 20-OCT-2003 (first entry)
 DT
 XX

DE Propionibacterium acnes predicted ORF-encoded polypeptide #12420.
KW Acne vulgaris; antisborrheic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

PN WO2003033515-A1.

XX 24-APR-2003.

PF 11-OCT-2002; 2002WO-US032727.

PR 15-OCT-2001; 2001US-00978825.

PA (CORI-) CORIXA CORP.

PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JI;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Valliave-Douglass J;

XX WPI: 2003-381789/36.

DR N-PSDB; ACF64480.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.

PS Example 1; SEQ ID NO 12420; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC encoding a Propionibacterium acnes protein. The invention also relates to
CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC immunogenic fragments of P. acnes polypeptides. The invention
CC additionally encompasses expression vectors and host cells comprising a
CC polynucleotide of the invention; antibodies against polypeptides of the
CC invention; fusion proteins comprising a polypeptide of the invention; a
CC method for stimulating an immune response specific for a P. acnes
CC polypeptide and an isolated T cell population comprising T cells prepared
CC via this method; a vaccine composition (comprising P. acnes polypeptides,
CC polynucleotides, antibodies, fusion proteins, T cell populations, or
CC antigen-presenting cells that express the polypeptide); a method and kit
CC for detecting or determining the presence or absence of P. acnes in a
CC patient; and a method for inhibiting the development of P. acnes in a
CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC proteins, T cell populations or antigen-presenting cells that express the
CC polypeptides are useful for diagnosing, preventing or treating acne
CC vulgaris, or for stimulating an immune response specific for a P. acnes
CC protein. The polynucleotides can also be used as probes or primers for
CC nucleic acid hybridisation. The vaccine composition is useful for the
CC stimulation of an immune response against P. acnes, or for treating acne,
CC and the kit is useful for performing a diagnostic assay. The present
CC sequence represents a polypeptide predicted to be encoded by an ORF (open
CC reading frame) contained within the P. acnes polynucleotides of the
CC invention. Note: The sequence data for this patent did not form part of
CC the printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 324 AA;

Query Match 100.0%; Score 36; DB 6; Length 324;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

DB 254 HPPHG 258
|||||

RESULT 40

ABU33219

ID ABU33219 standard; protein; 475 AA.

XX

AC ABU33219;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #18746.

DE Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Legionella pneumophila.

XX WO200277183-A2.

PN 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

PR 06-SEP-2001; 2001US-00948993.

PR 25-OCT-2001; 2001US-0342923P.

PR 08-FEB-2002; 2002US-00072851.

PR 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

PA Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI: 2003-023926/02.

DR N-PSDB; ACA37089.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

PS Claim 25; SEQ ID NO 61143; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than S. aureus, S. typhimurium,
CC K. pneumoniae or P. aeruginosa. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 475 AA;

Query Match 100.0%; Score 36; DB 6; Length 475;

Best Local Similarity 100.0%; Pred. No. 6.7e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5
|||
Db 412 HPPHG 416

Search completed: June 15, 2005, 14:15:50
Job time : 92 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

- 1: /cgn2_6/prodata/1/1aa/5A_COMB.pep:*
- 2: /cgn2_6/prodata/1/1aa/5B_COMB.pep:*
- 3: /cgn2_6/prodata/1/1aa/6A_COMB.pep:*
- 4: /cgn2_6/prodata/1/1aa/6B_COMB.pep:*
- 5: /cgn2_6/prodata/1/1aa/PCTUS_COMB.pep:*
- 6: /cgn2_6/prodata/1/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	36	100.0	231	4	US-09-257-583-9
2	36	100.0	262	4	US-09-489-039A-8374
3	36	100.0	480	4	US-09-252-991A-26799
4	36	100.0	516	4	US-09-252-991A-19788
5	36	100.0	613	4	US-09-252-991A-22424
6	36	100.0	1003	4	US-09-252-991A-24069
7	36	100.0	1447	3	US-09-041-886-25
8	36	100.0	1447	5	PCT-US94-05277-2
9	30	83.3	7	3	US-08-973-551-19
10	30	83.3	16	2	US-08-480-190-29
11	30	83.3	16	2	US-08-488-379-29
12	30	83.3	16	4	US-08-475-399A-29
13	30	83.3	16	4	US-08-077-255A-29
14	30	83.3	16	5	PCT-US93-07545-29
15	30	83.3	46	4	US-09-270-767-57079
16	30	83.3	60	4	US-09-248-796A-24991
17	30	83.3	64	4	US-09-248-796A-26393
18	30	83.3	69	4	US-09-107-433-2788
19	30	83.3	71	4	US-09-107-532A-5630
20	30	83.3	74	4	US-09-252-991A-24698
21	30	83.3	98	2	US-08-484-905-59
22	30	83.3	98	3	US-08-481-985B-59
23	30	83.3	98	3	US-08-370-476-59
24	30	83.3	109	4	US-09-513-999C-5250
25	30	83.3	119	1	US-08-469-486-50
26	30	83.3	119	2	US-08-469-658-50
27	30	83.3	139	4	US-09-621-978-6561

28	30	83.3	142	4	US-09-248-796A-22219	Sequence 22219, A
29	30	83.3	145	4	US-09-252-991A-25194	Sequence 25194, A
30	30	83.3	147	4	US-09-252-991A-16944	Sequence 16944, A
31	30	83.3	147	4	US-09-489-039A-9494	Sequence 9494, Ap
32	30	83.3	162	4	US-09-252-991A-22093	Sequence 22093, A
33	30	83.3	169	4	US-09-461-912A-47	Sequence 47, Appl
34	30	83.3	173	4	US-09-270-767-39233	Sequence 39233, A
35	30	83.3	173	4	US-09-270-767-54450	Sequence 54450, A
36	30	83.3	176	4	US-09-252-991A-24810	Sequence 24810, A
37	30	83.3	180	4	US-09-252-991A-28220	Sequence 28220, A
38	30	83.3	197	4	US-09-252-991A-20521	Sequence 20521, A
39	30	83.3	200	4	US-09-640-211A-743	Sequence 743, App
40	30	83.3	203	4	US-09-270-767-41835	Sequence 41835, A
41	30	83.3	216	1	US-08-451-947-8	Sequence 8, Appli
42	30	83.3	216	2	US-08-424-826A-8	Sequence 8, Appli
43	30	83.3	216	3	US-08-928-694-8	Sequence 8, Appli
44	30	83.3	216	4	US-08-450-842-8	Sequence 8, Appli
45	30	83.3	216	4	US-08-451-390-8	Sequence 8, Appli

ALIGNMENTS

RESULT 1

US-09-257-583-9

; Sequence 9, Application US/09257583A

; Patent No. 6429362

; GENERAL INFORMATION:

; APPLICANT: Crane, Virginia

; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters

; FILE REFERENCE: 5718-32, 035718/175219

; CURRENT APPLICATION NUMBER: US/09/257,583A

; CURRENT FILING DATE: 1999-02-25

; NUMBER OF SEQ ID NOS: 37

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 9

; LENGTH: 231

; TYPE: PRT

; ORGANISM: Zea mays

US-09-257-583-9

Query Match 100.0%; Score 36; DB 4; Length 231;

Best Local Similarity 100.0%; Pred. No. 52;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

Db 174 HPPHG 178

|||||

RESULT 2

US-09-489-039A-8374

; Sequence 8374, Application US/09489039A

; Patent No. 6610836

; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA

; FILE REFERENCE: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS

; CURRENT APPLICATION NUMBER: US/09/489,039A

; CURRENT FILING DATE: 2000-01-27

; PRIOR APPLICATION NUMBER: US 60/117,747

; PRIOR FILING DATE: 1999-01-29

; NUMBER OF SEQ ID NOS: 14342

; SEQ ID NO 8374

; LENGTH: 262

; TYPE: PRT

; ORGANISM: Klebsiella pneumoniae

US-09-489-039A-8374

Query Match 100.0%; Score 36; DB 4; Length 262;

Best Local Similarity 100.0%; Pred. No. 59;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      1 HPPHG 5
      |||||
Db      18 HPPHG 22

RESULT 3
US-09-252-991A-26799
; Sequence 26799, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 26799
; LENGTH: 480
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26799

Query Match      100.0%; Score 36; DB 4; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      358 HPPHG 362

RESULT 4
US-09-252-991A-19788
; Sequence 19788, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 19788
; LENGTH: 516
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-19788

Query Match      100.0%; Score 36; DB 4; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.le+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      63 HPPHG 67

RESULT 5
US-09-252-991A-22424
; Sequence 22424, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
```

```
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22424
; LENGTH: 613
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22424

Query Match      100.0%; Score 36; DB 4; Length 613;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      467 HPPHG 471

RESULT 6
US-09-252-991A-24069
; Sequence 24069, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24069
; LENGTH: 1003
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24069

Query Match      100.0%; Score 36; DB 4; Length 1003;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      232 HPPHG 236

RESULT 7
US-09-041-886-25
; Sequence 25, Application US/09041886
; Patent No. 6235872
; GENERAL INFORMATION:
; APPLICANT: Bredesen, Dale E.
; APPLICANT: Rabizadeh, Sharroz
; TITLE OF INVENTION: Proapoptotic Peptides, Dependence
; TITLE OF INVENTION: Polypeptides and Methods of Use
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Campbell & Flores LLP
; STREET: 4370 La Jolla Village Drive, Suite 700
; CITY: San Diego
; STATE: California
; COUNTRY: United States
```

ZIP: 92122
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/041,886
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Campbell, Cathryn A.
REGISTRATION NUMBER: 31,815
REFERENCE/DOCKET NUMBER: P-LJ 2626
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 535-9001
TELEFAX: (619) 535-8949
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-041-886-25

Query Match 100.0%; Score 36; DB 3; Length 1447;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1084 HPPHG 1088

QY 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 8
PCT-US94-05277-2
Sequence 2, Application PC/TUS9405277
GENERAL INFORMATION:
APPLICANT: Bruskin, Arthur
APPLICANT: Jarosz, David E.
APPLICANT: Johnson, Karen
APPLICANT: Kinzler, Kenneth W.
APPLICANT: Vogelstein, Bert
APPLICANT: Zabrecky, James R.
TITLE OF INVENTION: Antibodies Specific for DCC Gene Product
NUMBER OF SEQUENCES: 2
CORRESPONDENCE ADDRESS:
ADDRESSEE: Banner, Birch, McKie & Beckett
STREET: 1001 G Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US94/05277
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Kagan, Sarah A.
REGISTRATION NUMBER: 32,141
REFERENCE/DOCKET NUMBER: 01107.42709
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202.508.9100
TELEFAX: 202.508.9299
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1447 amino acids

TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US94-05277-2

Query Match 100.0%; Score 36; DB 5; Length 1447;
Best Local Similarity 100.0%; Pred. No. 3e+02; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 9
US-08-973-551-19
Sequence 19, Application US/08973551
Patent No. 6113902
GENERAL INFORMATION:
APPLICANT: Chermann, Jean-Claude
APPLICANT: Le Contel, Carole
APPLICANT: Galea, Pascale
TITLE OF INVENTION: VACCINE AGAINST INFECTIOUS AGENTS HAVING
AN INTRACELLULAR PHASE, COMPOSITION FOR THE TREATMENT AND
PREVENTION OF HIV INFECTIONS, ANTIBODIES AND METHOD OF
DIAGNOSIS
TITLE OF INVENTION: DIAGNOSIS
NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 3000 K Street, N.W., Suite 500
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/973,551
FILING DATE: 30-DEC-1997
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR96/01006
FILING DATE: 28-JUN-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 9507914
FILING DATE: 30-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: Granados, Patricia D.
REGISTRATION NUMBER: 33,683
REFERENCE/DOCKET NUMBER: 65691/130
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202)672-5300
TELEFAX: (202)672-5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 7 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-973-551-19

Query Match 83.3%; Score 30; DB 3; Length 7;
Best Local Similarity 100.0%; Pred. No. 4.1e+05;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 2 HPPH 5

```
RESULT 10
US-08-480-190-29
; Sequence 29, Application US/08480190
; Patent No. 5827516
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/480,190
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-480-190-29

Query Match      83.3%; Score 30; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      1 HPPH 4

RESULT 11
US-08-488-379-29
; Sequence 29, Application US/08488379
; Patent No. 5850103
; GENERAL INFORMATION:
; APPLICANT: Robert G. Urban
; APPLICANT: Roman M. Chicz
; APPLICANT: Dario A. A. Vignali
; APPLICANT: Mary L. Hedley
; APPLICANT: Lawrence J. Stern
; APPLICANT: Jack L. Strominger
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
```

```
; NUMBER OF SEQUENCES: 274
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: MS-DOS (Version 5.0)
; SOFTWARE: WordPerfect (Version 5.1)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,379
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/077,255
; FILING DATE: June 15, 1993
; APPLICATION NUMBER: 07/925,460
; FILING DATE: August 11, 1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Clark, Paul T.
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 00246/168001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; US-08-488-379-29

Query Match      83.3%; Score 30; DB 2; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
      ||||
Db      1 HPPH 4

RESULT 12
US-08-475-399A-29
; Sequence 29, Application US/08475399A
; Patent No. 6509033
; GENERAL INFORMATION:
; APPLICANT: Urban, Robert G.
; APPLICANT: Chicz, Roman M.
; APPLICANT: Vignali, Dario A.A.
; APPLICANT: Hedley, Mary L.
; APPLICANT: Stern, Lawrence J.
; APPLICANT: Strominger, Jack L.
; TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
; NUMBER OF SEQUENCES: 276
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: US
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

APPLICATION NUMBER: US/08/475,399A
FILING DATE: 07-JUN-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,255
FILING DATE: 15-JUN-1993
APPLICATION NUMBER: 07/925,460
FILING DATE: 11-AUG-1992
ATTORNEY/AGENT INFORMATION:
NAME: Fraser, Janis K.
REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00246/168003
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-507
TELEFAX: 617/542-890
TELEX: 200154
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 16 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-475-399A-29

Query Match 83.3%; Score 30; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 13
US-08-077-255A-29
Sequence 29, Application US/08077255A
Patent No. 6696061
GENERAL INFORMATION:
APPLICANT: Robert G. Urban
APPLICANT: Roman M. Chicz
APPLICANT: Dario A. A. Vignali
APPLICANT: Mary L. Hedley
APPLICANT: Lawrence J. Stern
APPLICANT: Jack L. Strominger
TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
NUMBER OF SEQUENCES: 274
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/077,255A
FILING DATE: June 15, 1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/925,460
FILING DATE: August 11, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00246/168001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:

LENGTH: 16
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
US-08-077-255A-29

Query Match 83.3%; Score 30; DB 4; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 14
PCT-US93-07545-29
Sequence 29, Application PC/TUS9307545
GENERAL INFORMATION:
APPLICANT: Robert G. Urban
APPLICANT: Roman M. Chicz
APPLICANT: Dario A. A. Vignali
APPLICANT: Mary L. Hedley
APPLICANT: Lawrence J. Stern
APPLICANT: Jack L. Strominger
TITLE OF INVENTION: IMMUNOMODULATORY PEPTIDES
NUMBER OF SEQUENCES: 273
CORRESPONDENCE ADDRESS:
ADDRESSEE: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: Massachusetts
COUNTRY: U.S.A.
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/07545
FILING DATE: 19930811
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/925,460
FILING DATE: August 11, 1992
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00246/168001
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 29:
SEQUENCE CHARACTERISTICS:
LENGTH: 16
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-07545-29

Query Match 83.3%; Score 30; DB 5; Length 16;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 1 HPPH 4

RESULT 15
US-09-270-767-57079

; Sequence 57079, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 57079
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-57079

Query Match 83.3%; Score 30; DB 4; Length 46;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|
|
|
|
Db 9 HPPH 12

RESULT 16
US-09-248-796A-24991
; Sequence 24991, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 24991
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-24991

Query Match 83.3%; Score 30; DB 4; Length 60;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|
|
|
|
Db 14 HPPH 17

RESULT 17
US-09-248-796A-26393
; Sequence 26393, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICAN
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 26393

; LENGTH: 64
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-26393

Query Match 83.3%; Score 30; DB 4; Length 64;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|
|
|
|
Db 47 HPPH 50

RESULT 18
US-09-107-433-2788
; Sequence 2788, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354

COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: <Unknown>
OPERATING SYSTEM: <Unknown>
SOFTWARE: <Unknown>

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,433
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085131
FILING DATE: May 12, 1998
APPLICATION NUMBER: 60/051553
FILING DATE: July 2, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneke
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-011
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277

INFORMATION FOR SEQ ID NO: 2788:
SEQUENCE CHARACTERISTICS:
LENGTH: 69 amino acids
TYPE: amino acid
TOPOLOGY: linear

MOLECULE TYPE: protein
HYPOTHETICAL: YES
ORIGINAL SOURCE:
ORGANISM: Streptococcus pneumoniae

FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...69
SEQUENCE DESCRIPTION: SEQ ID NO: 2788:

US-09-107-433-2788

Query Match 83.3%; Score 30; DB 4; Length 69;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|
|
|
|
Db 6 HPPH 9

```
RESULT 19
; US-09-107-532A-5630
; Sequence 5630, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESS: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5630:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (B) LOCATION 1...71
; SEQUENCE DESCRIPTION: SEQ ID NO: 5630:
US-09-107-532A-5630
Query Match      83.3%; Score 30; DB 4; Length 71;
Best Local Similarity 100.0%; Pred. No. 1.7e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY      1 HPPH 4
      ||||
DB      51 HPPH 54

RESULT 20
; US-09-252-991A-24698
; Sequence 24698, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.116
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
```

```
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24698
; LENGTH: 74
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24698
Query Match      83.3%; Score 30; DB 4; Length 74;
Best Local Similarity 100.0%; Pred. No. 1.8e+02; Indels 0; Gaps 0;
Matches 4; Conservative 0; Mismatches 0;

QY      1 HPPH 4
      ||||
DB      9 HPPH 12

RESULT 21
; US-08-484-905-59
; Sequence 59, Application US/08484905
; Patent No. 5976551
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Philippe
; TITLE OF INVENTION: An Altered Major Histocompatibility
; TITLE OF INVENTION: Complex(MHC) Determinant and Methods for Using the
; TITLE OF INVENTION: Determinant
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS-/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484,905
; FILING DATE: 07-JUNE-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Potter, Jane E. R.
; REGISTRATION NUMBER: 33,332
; REFERENCE/DOCKET NUMBER: 03495.0106-03000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-484-905-59
Query Match      83.3%; Score 30; DB 2; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
```

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|||
Db 31 HPPH 34

RESULT 22

US-08-481-985B-59
; Sequence 59, Application US/08481985B
; Patent No. 601146
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre
; APPLICANT: Kourilsky, Phillipe
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 148
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 03495.0106-04000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4400
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-481-985B-59

Query Match 83.3%; Score 30; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|||
Db 31 HPPH 34

RESULT 23

US-08-370-476-59
; Sequence 59, Application US/08370476
; Patent No. 6153408
; GENERAL INFORMATION:
; APPLICANT: Mottez, Estelle
; APPLICANT: Abastado, Jean-Pierre

; APPLICANT: Kourilsky, Phillipe
; APPLICANT: Lone, Yu-Chun
; APPLICANT: Ojcius, David
; APPLICANT: Casrouge, Armanda
; TITLE OF INVENTION: Altered Major Histocompatibility Complex
; TITLE OF INVENTION:
; NUMBER OF SEQUENCES: 127
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W., Suite 700
; CITY: Washington
; STATE: D.C.
; ZIP: 20005-3315
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,476
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/117,575
; FILING DATE: 07-SEP-1993
; APPLICATION NUMBER: US 08/072,787
; FILING DATE: 06-JUN-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/801,818
; FILING DATE: 05-DEC-1991
; APPLICATION NUMBER: US 07/792,473
; FILING DATE: 15-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyers, Kenneth J.
; REGISTRATION NUMBER: 25,146
; REFERENCE/DOCKET NUMBER: 05243.0001-01000
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-408-4000
; TELEFAX: 202-408-4400
; INFORMATION FOR SEQ ID NO: 59:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 98 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-370-476-59

Query Match 83.3%; Score 30; DB 3; Length 98;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPH 4
|||
Db 31 HPPH 34

RESULT 24

US-09-513-999C-5250
; Sequence 5250, Application US/09513999C
; Patent No. 6783961
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert, A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6783961
; FILE REFERENCE: 59.US2.REG
; CURRENT APPLICATION NUMBER: US/09/513,999C
; CURRENT FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/122,487
; PRIOR FILING DATE: 1999-02-26

; NUMBER OF SEQ ID NOS: 36681
; SOFTWARE: Patent.pm
; SEQ ID NO 5250
; LENGTH: 109
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-513-999C-5250

Query Match 83.3%; Score 30; DB 4; Length 109;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 72 HPPH 75

RESULT 25
US-08-469-486-50
; Sequence 50, Application US/08469486
; Patent No. 5739281
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Ezerodt, Michael
; TITLE OF INVENTION: Improved method for the refolding of
; TITLE OF INVENTION: proteins
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,486
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-486-50

Query Match 83.3%; Score 30; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 51 HPPH 54

RESULT 26
US-08-469-658-50
; Sequence 50, Application US/08469658
; Patent No. 5917018
; GENERAL INFORMATION:
; APPLICANT: Thøgersen, Hans Christian
; APPLICANT: Holtet, Thor Las
; APPLICANT: Ezerodt, Michael
; TITLE OF INVENTION: IMPROVED METHOD FOR THE REFOOLDING OF
; TITLE OF INVENTION: PROTEINS
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version
; SOFTWARE: #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,658
; FILING DATE: June 5, 1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/192,060
; FILING DATE: February 4, 1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Paul T. Clark
; REGISTRATION NUMBER: 30,162
; REFERENCE/DOCKET NUMBER: 06363/002002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617 542 5070
; TELEFAX: 617 542 8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 50:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 119 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-469-658-50

Query Match 83.3%; Score 30; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 51 HPPH 54

RESULT 27
US-09-621-976-6561
; Sequence 6561, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm

; SEQ ID NO 6561
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-621-976-6561

Query Match 83.3%; Score 30; DB 4; Length 139;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 51 HPPH 54

RESULT 28

US-09-248-796A-22219
; Sequence 22219, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:

; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22219
; LENGTH: 142
; TYPE: PRT
; ORGANISM: Candida albicans
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (31)
; OTHER INFORMATION: Identity of amino acid sequences at the above locations are unknown

US-09-248-796A-22219

Query Match 83.3%; Score 30; DB 4; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 11 HPPH 14

RESULT 29

US-09-252-991A-25194
; Sequence 25194, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 25194
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25194

Query Match 83.3%; Score 30; DB 4; Length 145;

Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 55 HPPH 58

RESULT 30

US-09-252-991A-16944
; Sequence 16944, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 16944
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-16944

Query Match 83.3%; Score 30; DB 4; Length 147;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 112 HPPH 115

RESULT 31

US-09-489-039A-9494
; Sequence 9494, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:

; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 9494
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-9494

Query Match 83.3%; Score 30; DB 4; Length 147;
Best Local Similarity 80.0%; Pred. No. 3.5e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 14 YPPHG 18

RESULT 32

US-09-252-991A-22093
; Sequence 22093, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:

```
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; PRIOR FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 22093
; LENGTH: 162
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-22093

Query Match      83.3%; Score 30; DB 4; Length 162;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 50 HPPH 53

RESULT 33
US-09-461-912A-47
; Sequence 47, Application US/09461912A
; Patent No. 6709855
; GENERAL INFORMATION:
; APPLICANT: Stanton, Lawrence A.
; APPLICANT: White, R. Tyler
; APPLICANT: Damm, Deborah L.
; APPLICANT: Lewicki, John A.
; TITLE OF INVENTION: Methods for detection and use of
; FILE REFERENCE: SCIOS.011A
; CURRENT APPLICATION NUMBER: US/09/461,912A
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/113,008
; PRIOR FILING DATE: 1998-12-18
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 47
; LENGTH: 169
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-461-912A-47

Query Match      83.3%; Score 30; DB 4; Length 169;
Best Local Similarity 100.0%; Pred. No. 3.9e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 43 HPPH 46

RESULT 34
US-09-270-767-39233
; Sequence 39233, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 39233
; LENGTH: 173
```

```
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-39233

Query Match      83.3%; Score 30; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 76 HPPH 79

RESULT 35
US-09-270-767-54450
; Sequence 54450, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 54450
; LENGTH: 173
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-54450

Query Match      83.3%; Score 30; DB 4; Length 173;
Best Local Similarity 100.0%; Pred. No. 4e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 76 HPPH 79

RESULT 36
US-09-252-991A-24810
; Sequence 24810, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 24810
; LENGTH: 176
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-24810

Query Match      83.3%; Score 30; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 4.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 99 HPPH 102
```

```
RESULT 37
US-09-252-991A-28220
; Sequence 28220, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 28220
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-28220

Query Match      83.3%; Score 30; DB 4; Length 180;
Best Local Similarity 100.0%; Pred. No. 4.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      133 HPPH 136

RESULT 38
US-09-252-991A-20521
; Sequence 20521, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 20521
; LENGTH: 197
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-20521

Query Match      83.3%; Score 30; DB 4; Length 197;
Best Local Similarity 80.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      4 YPPHG 8

RESULT 39
US-09-640-211A-743
; Sequence 743, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
```

```
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 743
; LENGTH: 200
; TYPE: PRT
; ORGANISM: Eucalyptus grandis
US-09-640-211A-743

Query Match      83.3%; Score 30; DB 4; Length 200;
Best Local Similarity 100.0%; Pred. No. 4.6e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      72 HPPH 75

RESULT 40
US-09-270-767-41835
; Sequence 41835, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 41835
; LENGTH: 203
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-41835

Query Match      83.3%; Score 30; DB 4; Length 203;
Best Local Similarity 100.0%; Pred. No. 4.7e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPH 4
Db      9 HPPH 12

Search completed: June 15, 2005, 14:24:20
Job time : 24.5 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	5	14 US-10-074-225A-9	Sequence 9, Appli
2	36	100.0	20	9 US-09-864-761-41108	Sequence 41108, A
3	36	100.0	53	16 US-10-425-115-338760	Sequence 338760,
4	36	100.0	63	9 US-09-764-869-833	Sequence 833, App
5	36	100.0	63	14 US-10-091-504-833	Sequence 833, App
6	36	100.0	63	15 US-10-227-577-833	Sequence 833, App
7	36	100.0	76	15 US-10-424-599-200189	Sequence 200189,
8	36	100.0	92	16 US-10-425-115-345170	Sequence 345170,
9	36	100.0	101	14 US-10-074-225A-6	Sequence 6, Appli
10	36	100.0	112	15 US-10-424-599-149766	Sequence 149766,
11	36	100.0	117	15 US-10-466-164-37	Sequence 37, Appli

12	36	100.0	118	14	US-10-023-896-137	Sequence 137, App
13	36	100.0	123	16	US-10-437-963-179511	Sequence 179511,
14	36	100.0	126	16	US-10-425-115-361744	Sequence 261744,
15	36	100.0	135	15	US-10-425-114-42907	Sequence 42907, A
16	36	100.0	150	16	US-10-425-115-293118	Sequence 293118,
17	36	100.0	156	16	US-10-425-115-249025	Sequence 249025,
18	36	100.0	168	16	US-10-425-115-265234	Sequence 265234,
19	36	100.0	170	16	US-10-425-115-251226	Sequence 251226,
20	36	100.0	181	15	US-10-424-599-222089	Sequence 222089,
21	36	100.0	230	16	US-10-425-115-317091	Sequence 317091,
22	36	100.0	231	9	US-09-840-479-9	Sequence 9, Appli
23	36	100.0	240	16	US-10-437-963-117204	Sequence 117204,
24	36	100.0	260	15	US-10-425-114-64276	Sequence 64276, A
25	36	100.0	262	11	US-09-764-875-518	Sequence 918, App
26	36	100.0	264	15	US-10-425-114-64304	Sequence 64304, A
27	36	100.0	265	16	US-10-425-115-218724	Sequence 218724,
28	36	100.0	313	16	US-10-437-963-155484	Sequence 155484,
29	36	100.0	346	16	US-10-437-963-173538	Sequence 173538,
30	36	100.0	426	15	US-10-425-114-70512	Sequence 70512, A
31	36	100.0	475	15	US-10-282-122A-61143	Sequence 61143, A
32	36	100.0	517	16	US-10-437-963-106725	Sequence 106725,
33	36	100.0	526	14	US-10-074-225A-3	Sequence 3, Appli
34	36	100.0	1498	16	US-10-437-963-200249	Sequence 200249,
35	30	83.3	7	10	US-09-827-345-19	Sequence 19, Appli
36	30	83.3	15	15	US-10-417-895A-39	Sequence 39, Appli
37	30	83.3	25	9	US-09-864-761-40555	Sequence 40555, A
38	30	83.3	25	9	US-09-864-761-47241	Sequence 47241, A
39	30	83.3	38	14	US-10-092-750-46	Sequence 46, Appli
40	30	83.3	38	16	US-10-425-115-263123	Sequence 263123,
41	30	83.3	40	16	US-10-767-701-47944	Sequence 47944, A
42	30	83.3	41	15	US-10-424-599-159525	Sequence 159525,
43	30	83.3	41	15	US-10-424-599-177264	Sequence 177264,
44	30	83.3	43	10	US-09-813-153-129	Sequence 129, App
45	30	83.3	43	15	US-10-415-383-9	Sequence 9, Appli

ALIGNMENTS

RESULT 1

US-10-074-225A-9
; Sequence 9, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; TITLE OF INVENTION: ANTI-TUMOR AGENT
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 9
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-9

Query Match 100.0%; Score 36; DB 14; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.5e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 HPPHG 5

Db 1 HPPHG 5

RESULT 2
US-09-864-761-41108
; Sequence 41108, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aeonica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; PRIOR FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 41108
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC026236.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 2.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.86
US-09-864-761-41108
Query Match 100.0%; Score 36; DB 9; Length 20;
Best Local Similarity 100.0%; Pred. No. 42;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3 HPPHG 7
RESULT 3
US-10-425-115-338760
; Sequence 338760, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 338760
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_72114C.1.pgp
US-10-425-115-338760
Query Match 100.0%; Score 36; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 96;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 27 HPPHG 31
RESULT 4
US-09-764-869-833
; Sequence 833, Application US/09764869
; Patent No. US20020061521A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007
; CURRENT APPLICATION NUMBER: US/09/764,869
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-869-833
Query Match 100.0%; Score 36; DB 9; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 26 HPPHG 30

```
RESULT 5
US-10-091-504-833
; Sequence 833, Application US/10091504
; Publication No. US20030059908A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C1
; CURRENT APPLICATION NUMBER: US/10/091,504
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 2442
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-091-504-833

Query Match          100.0%; Score 36; DB 14; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      26 HPPHG 30

RESULT 6
US-10-227-577-833
; Sequence 833, Application US/10227577
; Publication No. US20040005575A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC007C2
; CURRENT APPLICATION NUMBER: US/10/227,577
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/091,504
; PRIOR FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: 09/764,869
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: 60/214,886
; PRIOR FILING DATE: 2000-06-28
; PRIOR APPLICATION NUMBER: 60/217,487
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,758
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/220,963
; PRIOR FILING DATE: 2000-07-26
; PRIOR APPLICATION NUMBER: 60/217,496
; PRIOR FILING DATE: 2000-07-11
; PRIOR APPLICATION NUMBER: 60/225,447
; PRIOR FILING DATE: 2000-08-14
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 2442
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 833

Query Match          100.0%; Score 36; DB 14; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      26 HPPHG 30

RESULT 7
US-10-424-599-200189
; Sequence 200189, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200189
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22798C.1.pap
; US-10-424-599-200189

Query Match          100.0%; Score 36; DB 15; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      47 HPPHG 51

RESULT 8
US-10-425-115-345170
; Sequence 345170, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (50)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (55)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (58)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; US-10-227-577-833

Query Match          100.0%; Score 36; DB 15; Length 63;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      26 HPPHG 30
```

```
RESULT 7
US-10-424-599-200189
; Sequence 200189, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 200189
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(76)
; OTHER INFORMATION: unsure at all Xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_22798C.1.pap
; US-10-424-599-200189
```

```
Query Match          100.0%; Score 36; DB 15; Length 76;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      47 HPPHG 51
```

```
RESULT 8
US-10-425-115-345170
; Sequence 345170, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
```

```

; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 345170
; LENGTH: 92
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77962C.1.pap
US-10-425-115-345170

Query Match      100.0%; Score 36; DB 16; Length 92;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      59 HPPHG 63

RESULT 9
US-10-074-225A-6
; Sequence 6, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 101
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-6

Query Match      100.0%; Score 36; DB 14; Length 101;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      33 HPPHG 37

RESULT 10
US-10-424-599-149766
; Sequence 149766, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 149766
; LENGTH: 112
; TYPE: PRT
; ORGANISM: Glycine max

```

```

; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_106260C.1.pap
US-10-424-599-149766

Query Match      100.0%; Score 36; DB 15; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      18 HPPHG 22

RESULT 11
US-10-466-164-37
; Sequence 37, Application US/10466164
; Publication No. US20040058365A1
; GENERAL INFORMATION:
; APPLICANT: INCYTE CORPORATION; PANZER, Scott R.;
; APPLICANT: LINCOLN, Stephen E.; ALTUS, Christina M.;
; APPLICANT: DUFOUR, Gerard E.; JACKSON, Jennifer L.;
; APPLICANT: JONES, Anissa L.; DAM, Tam C.;
; APPLICANT: LIU, Tommy F.; HARRIS, Bernard;
; APPLICANT: FLORES, Vincent Z.; DAFPO, Abel;
; APPLICANT: MARWAHA, Rakesh; CHEN, Alice J.;
; APPLICANT: CHANG, Simon C.; GERSTIN, Jr., Edward H.;
; APPLICANT: PERALTA, Careyna H.; DAVID, Marie H.;
; APPLICANT: LEWIS, Samantha A.
; TITLE OF INVENTION: MOLECULES FOR DISEASE DETECTION AND TREATMENT
; FILE REFERENCE: PT-1215 PCT
; CURRENT APPLICATION NUMBER: US/10/466,164
; CURRENT FILING DATE: 2003-07-11
; PRIOR APPLICATION NUMBER: PCT/US02/01008
; PRIOR FILING DATE: 2002-01-09
; PRIOR APPLICATION NUMBER: US 60/261,865
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/263,065
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/263,329
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/262,209
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,208
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/262,326
; PRIOR FILING DATE: 2001-01-17
; PRIOR APPLICATION NUMBER: US 60/263,063
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: US 60/261,622
; NUMBER OF SEQ ID NOS: 72
; SOFTWARE: PERL Program
; SEQ ID NO 37
; LENGTH: 117
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20040058365A1 LI:180252.16.orf2:2001JAN12
US-10-466-164-37

Query Match      100.0%; Score 36; DB 15; Length 117;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      17 HPPHG 21

RESULT 12
US-10-023-896-137
; Sequence 137, Application US/10023896

```

```
; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Victor Roschke
; TITLE OF INVENTION: 29 Human Cancer Associated Proteins
; FILE REFERENCE: PA004P1
; CURRENT APPLICATION NUMBER: US/10/023,896
; CURRENT FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: unassigned
; PRIOR FILING DATE: 2001-12-21
; PRIOR APPLICATION NUMBER: PCT/US00/23794
; PRIOR FILING DATE: 2000-08-30
; PRIOR APPLICATION NUMBER: 60/152,296
; PRIOR FILING DATE: 1999-09-03
; PRIOR APPLICATION NUMBER: 60/158,003
; PRIOR FILING DATE: 1999-10-06
; NUMBER OF SEQ ID NOS: 138
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 137
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-023-896--137

Query Match          100.0%; Score 36; DB 14; Length 118;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 96 HPPHG 100

RESULT 13
US-10-437-963-179511
; Sequence 179511, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 179511
; LENGTH: 123
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_76967C.1.pep
US-10-437-963-179511

Query Match          100.0%; Score 36; DB 16; Length 123;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 49 HPPHG 53

RESULT 14
US-10-425-115-261744
; Sequence 261744, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
US-10-425-115-261744

; Publication No. US20030027776A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261744
; LENGTH: 126
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170323C.1.pep
US-10-425-115-261744

Query Match          100.0%; Score 36; DB 16; Length 126;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 85 HPPHG 89

RESULT 15
US-10-425-114-42907
; Sequence 42907, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 42907
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700570105_FLI.pep
US-10-425-114-42907

Query Match          100.0%; Score 36; DB 15; Length 135;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 108 HPPHG 112

RESULT 16
US-10-425-115-293118
; Sequence 293118, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: Plants and Uses Thereof for Plant Improvement
US-10-425-115-293118
```

; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 293118
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(150)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_30403C.1.pep
US-10-425-115-293118

Query Match 100.0%; Score 36; DB 16; Length 150;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 5 HPPHG 9

RESULT 17

US-10-425-115-249025
; Sequence 249025, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 249025
; LENGTH: 156
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_158698C.1.pep
US-10-425-115-249025

Query Match 100.0%; Score 36; DB 16; Length 156;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 51 HPPHG 55

RESULT 18

US-10-425-115-265234
; Sequence 265234, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 265234
; LENGTH: 168

; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_1734C.1.pep
US-10-425-115-265234

Query Match 100.0%; Score 36; DB 16; Length 168;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 58 HPPHG 62

RESULT 19

US-10-425-115-251226
; Sequence 251226, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 251226
; LENGTH: 170
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(170)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_160703C.1.pep
US-10-425-115-251226

Query Match 100.0%; Score 36; DB 16; Length 170;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 149 HPPHG 153

RESULT 20

US-10-424-599-222089
; Sequence 222089, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 222089
; LENGTH: 181
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_42576C.1.pep
US-10-424-599-222089

```
Query Match      100.0%; Score 36; DB 15; Length 181;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      23 HPPHG 27

RESULT 21
US-10-425-115-317091
; Sequence 317091, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 317091
; LENGTH: 230
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(230)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_52259C.1.pep
US-10-425-115-317091

Query Match      100.0%; Score 36; DB 16; Length 230;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      170 HPPHG 174

RESULT 22
US-09-840-479-9
; Sequence 9, Application US/09840479
; Patent No. US20010025380A1
; GENERAL INFORMATION:
; APPLICANT: Crane, Virginia
; TITLE OF INVENTION: Family Of Maize PR-1 Genes And Promoters
; FILE REFERENCE: 5718-32, 035718/175219
; CURRENT APPLICATION NUMBER: US/09/840,479
; CURRENT FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: 09/257,583
; PRIOR FILING DATE: 1999-02-25
; NUMBER OF SEQ ID NOS: 37
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 9
; LENGTH: 231
; TYPE: PRT
; ORGANISM: Zea mays
US-09-840-479-9

Query Match      100.0%; Score 36; DB 9; Length 231;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      174 HPPHG 178
```

```
RESULT 23
US-10-437-963-117204
; Sequence 117204, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 117204
; LENGTH: 240
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_20631C.1.pep
US-10-437-963-117204

Query Match      100.0%; Score 36; DB 16; Length 240;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      216 HPPHG 220

RESULT 24
US-10-425-114-64276
; Sequence 64276, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64276
; LENGTH: 260
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3607-056-B7_FLI.pep
US-10-425-114-64276

Query Match      100.0%; Score 36; DB 15; Length 260;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      203 HPPHG 207

RESULT 25
```

```
US-09-764-875-918
; Sequence 918, Application US/09764875
; Publication No. US20040018969A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: FU202
; CURRENT APPLICATION NUMBER: US/09/764,875
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1249
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 918
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-764-875-918

Query Match          100.0%; Score 36; DB 11; Length 262;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      17 HPPHG 21

RESULT 26
US-10-425-114-64304
; Sequence 64304, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64304
; LENGTH: 264
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3608-049-C7_FLI.pep
US-10-425-114-64304

Query Match          100.0%; Score 36; DB 15; Length 264;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      207 HPPHG 211

RESULT 27
US-10-425-115-218724
; Sequence 218724, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
```

```
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 218724
; LENGTH: 265
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_131068C.1.pep
US-10-425-115-218724

Query Match          100.0%; Score 36; DB 16; Length 265;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      207 HPPHG 211

RESULT 28
US-10-437-963-155484
; Sequence 155484, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155484
; LENGTH: 313
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55242C.1.pep
US-10-437-963-155484

Query Match          100.0%; Score 36; DB 16; Length 313;
Best Local Similarity 100.0%; Pred. No. 4.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
DB      178 HPPHG 182

RESULT 29
US-10-437-963-173538
; Sequence 173538, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
```

; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 173538
; LENGTH: 346
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_71566C.1.pap
US-10-437-963-173538

Query Match 100.0%; Score 36; DB 16; Length 346;
Best Local Similarity 100.0%; Pred. No. 4.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
DB 5 HPPHG 9

RESULT 30
US-10-425-114-70512
; Sequence 70512, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaaka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70512
; LENGTH: 426
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-2MFLM017050F06_FLI.pap
US-10-425-114-70512

Query Match 100.0%; Score 36; DB 15; Length 426;
Best Local Similarity 100.0%; Pred. No. 5.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
DB 19 HPPHG 23

RESULT 31
US-10-282-122A-61143
; Sequence 61143, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITPA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A

; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 61143
; LENGTH: 475
; TYPE: PRT
; ORGANISM: Legionella pneumophila
US-10-282-122A-61143

Query Match 100.0%; Score 36; DB 15; Length 475;
Best Local Similarity 100.0%; Pred. No. 6.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
DB 412 HPPHG 416

RESULT 32
US-10-437-963-106725
; Sequence 106725, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 106725
; LENGTH: 517
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_11143C.1.pap
US-10-437-963-106725

Query Match 100.0%; Score 36; DB 16; Length 517;
Best Local Similarity 100.0%; Pred. No. 6.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
DB 382 HPPHG 386

RESULT 33
US-10-074-225A-3
; Sequence 3, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC
; FILE REFERENCE: 38342-178463
; CURRENT APPLICATION NUMBER: US/10/074,225A
; CURRENT FILING DATE: 2002-02-14
; PRIOR APPLICATION NUMBER: US 60/268,370
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3
; LENGTH: 526
; TYPE: PRT
; ORGANISM: Lepus americanus
US-10-074-225A-3

Query Match 100.0%; Score 36; DB 14; Length 526;
Best Local Similarity 100.0%; Pred. No. 6.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

QY 1 HPPHG 5
Db 353 HPPHG 357

RESULT 34
US-10-437-963-200249
; Sequence 200249, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 200249
; LENGTH: 1498
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_95737C.1.pap
US-10-437-963-200249

Query Match 100.0%; Score 36; DB 16; Length 1498;
Best Local Similarity 100.0%; Pred. No. 1.7e+03; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

QY 1 HPPHG 5
Db 1337 HPPHG 1341

RESULT 35
US-09-827-345-19

; Sequence 19, Application US/09827345
; Publication No. US20030021800A1
; GENERAL INFORMATION:
; APPLICANT: CHERMANN, JEAN-CLAUDE
; APPLICANT: LE CONTEL, CAROLE
; APPLICANT: GALEA, PASCALE
; TITLE OF INVENTION: VACCINE AGAINST INFECTIOUS AGENTS HAVING AN
; TITLE OF INVENTION: INTRACELLULAR PHASE, COMPOSITION FOR THE TREATMENT AND
; TITLE OF INVENTION: PREVENTION OF HIV INFECTIONS, ANTIBODIES AND METHOD OF
; TITLE OF INVENTION: DIAGNOSIS
; FILE REFERENCE: 065691-0216
; CURRENT APPLICATION NUMBER: US/09/827,345
; CURRENT FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: 09/599,549
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: PCT/FR96/01006
; PRIOR FILING DATE: 1996-06-28
; PRIOR APPLICATION NUMBER: 08/973,551
; PRIOR FILING DATE: 1998-02-19
; PRIOR APPLICATION NUMBER: FR 95/07914
; PRIOR FILING DATE: 1995-06-30
; NUMBER OF SEQ ID NOS: 28
; SOFTWARE: Patent in ver. 2.1
; SEQ ID NO 19
; LENGTH: 7
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-827-345-19

Query Match 83.3%; Score 30; DB 10; Length 7;
Best Local Similarity 100.0%; Pred. No. 1.5e+06; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 1 HPPH 4
Db 2 HPPH 5

RESULT 36
US-10-417-895A-39
; Sequence 39, Application US/10417895A
; Publication No. US20040033569A1
; GENERAL INFORMATION:
; APPLICANT: Crea, Roberto
; APPLICANT: Cappuccilli, Guido
; TITLE OF INVENTION: "DOPING" IN WALK-THROUGH MUTAGENESIS
; FILE REFERENCE: 1551.2002-001
; CURRENT APPLICATION NUMBER: US/10/417,895A
; CURRENT FILING DATE: 2003-04-16
; PRIOR APPLICATION NUMBER: 60/373,686
; PRIOR FILING DATE: 2002-04-17
; NUMBER OF SEQ ID NOS: 86
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 15
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: variant peptide for second complementarity
; OTHER INFORMATION: determining region of Fv region of an
; OTHER INFORMATION: immunoglobulin
US-10-417-895A-39

Query Match 83.3%; Score 30; DB 15; Length 15;
Best Local Similarity 100.0%; Pred. No. 2.9e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 4; Conservative 0;

QY 1 HPPH 4
Db 4 HPPH 7

RESULT 37
US-09-864-761-40555
; Sequence 40555, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 40555
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC007981.43
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 2.8
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.6
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 2.9
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 3.8
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 2.6
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3
; OTHER INFORMATION: EST_HUMAN HIT: A1026655.1, EVALUATE 3.60e-02
US-09-864-761-40555

Query Match 83.3%; Score 30; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
Db 10 HPPH 13

RESULT 38
US-09-864-761-47241
; Sequence 47241, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 47241
; LENGTH: 25
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC005326.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 0.47
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 0.41
; OTHER INFORMATION: EST_HUMAN HIT: H23753.1, EVALUATE 8.00e+00
US-09-864-761-47241

Query Match 83.3%; Score 30; DB 9; Length 25;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 19 HPPH 22

RESULT 39

US-10-092-750-46
; Sequence 46, Application US/10092750
; Publication No. US20030032157A1
; GENERAL INFORMATION:
; APPLICANT: Hammond, Philip W.
; APPLICANT: Alpin, Julia
; APPLICANT: Wright, Martin C.
; TITLE OF INVENTION: Polypeptides Interactive with BCL-XL
; FILE REFERENCE: 50036/050002
; CURRENT APPLICATION NUMBER: US/10/092,750
; CURRENT FILING DATE: 2002-03-07
; PRIOR APPLICATION NUMBER: US 60/274,526
; PRIOR FILING DATE: 2001-03-08
; NUMBER OF SEQ ID NOS: 253
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-092-750-46

Query Match 83.3%; Score 30; DB 14; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 34 HPPH 37

RESULT 40

US-10-425-115-263123
; Sequence 263123, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 263123
; LENGTH: 38
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_171580C.1.pep
US-10-425-115-263123

Query Match 83.3%; Score 30; DB 16; Length 38;
Best Local Similarity 100.0%; Pred. No. 6.5e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||
Db 28 HPPH 31

Search completed: June 15, 2005, 14:50:19
Job time : 80.75 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*
1: Pirl:*
2: Pirl:*
3: Pirl:*
4: Pirl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	118	2 F72674	hypothetical prote
2	36	100.0	176	2 F82478	hypothetical prote
3	36	100.0	235	2 T06168	pathogenesis relat
4	36	100.0	251	2 G72454	probable phosphoe
5	36	100.0	470	2 T46814	gamma-aminobutyrat
6	36	100.0	470	2 B95419	diaminobutyrat-py
7	36	100.0	1427	2 I51669	tumor suppressor -
8	36	100.0	1447	2 A54100	early E3 6.4K prot
9	30	83.3	61	1 ERAD64	hypothetical prote
10	30	83.3	73	2 H71311	hypothetical prote
11	30	83.3	107	2 T47876	hypothetical prote
12	30	83.3	112	2 A03861	hypothetical prote
13	30	83.3	113	2 B72629	hypothetical prote
14	30	83.3	113	2 S66936	probable membrane
15	30	83.3	115	2 T17301	hypothetical prote
16	30	83.3	119	1 MGMSB2	beta-2-microglobul
17	30	83.3	121	2 B72546	hypothetical prote
18	30	83.3	122	2 T46369	hypothetical prote
19	30	83.3	125	2 A05160	hypothetical prote
20	30	83.3	126	2 B93320	conserved hypotet
21	30	83.3	141	2 F71070	hypothetical prote
22	30	83.3	148	2 S39206	proline-rich prote
23	30	83.3	157	2 A99226	conserved hypotet
24	30	83.3	159	2 C84299	hypothetical prote
25	30	83.3	167	2 A97760	NADH2 dehydrogenas
26	30	83.3	169	2 JE0149	peripheral benzod
27	30	83.3	169	2 I38724	mitochondrial benz
28	30	83.3	169	2 I38105	peripheral benzodi
29	30	83.3	169	2 S14257	benzodiazepine rec

30	30	83.3	173	2	T51469	glycine/proline-ri
31	30	83.3	181	2	T52409	uclacyanin II [imp
32	30	83.3	183	2	B97784	formylmethionine d
33	30	83.3	183	2	AC2903	conserved hypotet
34	30	83.3	203	2	H69937	hypothetical prote
35	30	83.3	207	2	B84315	hypothetical prote
36	30	83.3	208	2	D75004	hypothetical prote
37	30	83.3	209	2	C89005	protein T24A6.3 [i
38	30	83.3	214	2	AF2901	conserved hypotet
39	30	83.3	214	2	H97676	hypothetical prote
40	30	83.3	216	2	I53100	eHAND - mouse
41	30	83.3	220	2	AH3444	calf thymus ribonu
42	30	83.3	221	2	S25168	CHT28 protein - ch
43	30	83.3	223	2	C87277	conserved hypotet
44	30	83.3	228	2	T02555	hypothetical prote
45	30	83.3	232	2	G72508	hypothetical prote

ALIGNMENTS

RESULT 1

F72674

hypothetical protein APE0819 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: F72674

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kawa
DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: F72674

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-118 <KAW>

A:Cross-references: UNIPROT:Q9YDU9; DDBJ:AP000060; NID:g5104188; PIDN:BAA79798.1; PID:d1

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE0819

C:Superfamily: Aeropyrum pernix hypothetical protein APE0819

Query Match 100.0%; Score 36; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5

|||||

Db 83 HPPHG 87

RESULT 2

F82478

hypothetical protein VCA0284 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #text_change 09-Jul-2004

C:Accession: F82478

R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;
chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragoi, I.; Sellers, P.
l. R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.

Nature 406, 477-483, 2000

A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.

A:Reference number: A82035; MUID:20406933; PMID:10952301

A:Accession: F82478

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-176 <HEI>

A:Cross-references: UNIPROT:Q9KMP0; GB:AE003853; NID:g9657677; PIDN:AAF9619;

A:Experimental source: serogroup O1; strain N16961; biotype El Tor

C:Genetics:

A:Gene: VCA0284

A:Map position: 2

```

Query Match      100.0%; Score 36; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      51 HPPHG 55

RESULT 3
T06168
pathogenesis related protein - barley
C;Species: Hordeum vulgare (barley)
C;Date: 30-Apr-1999 #sequence_revision 30-Apr-1999 #text_change 09-Jul-2004
C;Accession: T06168
R;Scott, K.J.
submitted to the EMBL Data Library, June 1994
A;Reference number: Z15503
A;Accession: T06168
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-235 <SCO>
A;Cross-references: UNIPROT:P16273; EMBL:X16648; NID:g499073; PIDN:CAA34641.1; PID:g4990
A;Experimental source: cv. Psakon 4

Query Match      100.0%; Score 36; DB 2; Length 235;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      176 HPPHG 180

RESULT 4
G72454
probable phosphoesterase (EC 3.1.1.-) APE2283 [similarity] - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: G72454
R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: G72454
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-251 <RAW>
A;Cross-references: UNIPROT:Q9Y9K5; DDBJ:AP000064; NID:G5105945; PIDN:BAA81295.1; PID:d1
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE2283
C;Superfamily: probable phosphoesterase MTH1882; phosphoesterase core homology

Query Match      100.0%; Score 36; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      172 HPPHG 176

RESULT 5
T46814
gamma-aminobutyrate aminotransferase (EC 2.6.1.-) rhbA [imported] - Rhizobium meliloti
C;Species: Rhizobium meliloti
C;Date: 17-Mar-2000 #sequence_revision 17-Mar-2000 #text_change 09-Jul-2004
C;Accession: T46814
R;Lynch, D.; O Connell, M.; O'Brien, J.
submitted to the EMBL Data Library, December 1998
A;Description: Cloning and sequence analysis of the Sinorhizobium meliloti 2011 rhizobad

```

```

A;Reference number: Z24097
A;Accession: T46814
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-470 <LYN>
A;Cross-references: UNIPROT:Q923R2; EMBL:AF110737; PIDN:AAD09412.1
A;Experimental source: strain 2011
C;Genetics:
A;Gene: rhbA
C;Function:
A;Pathway: siderophore biosynthesis
C;Superfamily: beta-alanine-pyruvate transaminase
C;Keywords: aminotransferase

Query Match      100.0%; Score 36; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      406 HPPHG 410

RESULT 6
B95419
diaminobutyrate-pyruvate transaminase (EC 2.6.1.46) RhbA [imported] - Sinorhizobium meli
C;Species: Sinorhizobium meliloti
C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
C;Accession: B95419
R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bowse
.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.;
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
A;Reference number: A95262; MUID:21396509; PMID:11481432
A;Accession: B95419
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-470 <KUR>
A;Cross-references: UNIPROT:Q923R2; GB:AE006469; PIDN:AAK65916.1; PID:g14524429; GSPDB:G
A;Experimental source: strain 1021, megaplasmid pSymA
R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;
L.; Hyman, R.W.; Jones, T.
Science 293, 668-672, 2001
A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C.
A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
A;Reference number: A96039; MUID:21368234; PMID:11474104
A;Contents: annotation
C;Genetics:
A;Gene: rhbA
C;Superfamily: beta-alanine-pyruvate transaminase
C;Keywords: aminotransferase

Query Match      100.0%; Score 36; DB 2; Length 470;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
      |||||
Db      406 HPPHG 410

RESULT 7
I51669
tumor suppressor - African clawed frog
C;Species: Xenopus laevis (African clawed frog)
C;Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 09-Jul-2004
C;Accession: I51669
R;Pierceall, W.E.; Reale, M.A.; Candia, A.F.; Wright, C.V.; Cho, K.R.; Fearon, E.R.
Dev. Biol. 166, 654-665, 1994
A;Title: Expression of a homologue of the deleted in colorectal cancer (DCC) gene in the

```

A;Reference number: I51668; MUID:95113183; PMID:7813784
A;Accession: I51669
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-1427 <PIE>
A;Cross-references: UNIPROT:Q91562; EMBL:U10986; NID:g606873; PIDN:AAA70168.1; PID:g606873
C;Genetics:
A;Gene: XDCCA

Query Match 100.0%; Score 36; DB 2; Length 1427;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1066 HPPHG 1070

RESULT 8
A54100
tumor suppressor protein DCC precursor - human
N;Alternate names: colorectal cancer suppressor DCC
C;Species: Homo sapiens (man)
C;Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C;Accession: A54100; A40098
R;Hedrick, L.; Cho, K.R.; Fearon, E.R.; Wu, T.C.; Kinzler, K.W.; Vogelstein, B.
Genes Dev. 8, 1174-1183, 1994
A;Title: The DCC gene product in cellular differentiation and colorectal tumorigenesis.
A;Reference number: A54100; MUID:95011532; PMID:7926722
A;Accession: A54100
A;Molecule type: mRNA
A;Residues: 1-1447 <HED>
R;Fearon, E.R.; Cho, K.R.; Nigro, J.M.; Kern, S.E.; Simons, J.W.; Ruppert, J.M.; Hamilton, S.C.
Science 247, 49-56, 1990
A;Title: Identification of a chromosome 18q gene that is altered in colorectal cancers.
A;Reference number: A40098; MUID:90100559; PMID:2294591
A;Accession: A40098
A;Molecule type: mRNA
A;Residues: 1-750 <FEA>
A;Cross-references: GB:M32292; NID:gi81492; PIDN:AAA35751.1; PID:gi81493
C;Genetics:
A;Gene: GDB:DCC
A;Cross-references: GDB:119838; OMIM:120470
A;Map position: 18q21.1-18q21.1
C;Keywords: transmembrane protein; tumor suppressor
F;1-25/Domain: signal sequence #status predicted <SIG>
F;26-1447/Product: tumor suppressor protein DCC #status predicted <MAT>

Query Match 100.0%; Score 36; DB 2; Length 1447;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 1084 HPPHG 1088

RESULT 9
ERAD64
early E3 6.4K protein - human adenovirus 35
C;Species: Mastadenovirus h35 (human adenovirus 35)
A;Note: host Homo sapiens (man)
C;Date: 31-Mar-1990 #sequence_revision 31-Mar-1990 #text_change 09-Jul-2004
C;Accession: B31162
R;Flomenberg, P.R.; Chen, M.; Horwitz, M.S.
J. Virol. 62, 4431-4437, 1988
A;Title: Sequence and genetic organization of adenovirus type 35 early region 3.
A;Reference number: A93039; MUID:89012230; PMID:3172347
A;Accession: B31162
A;Molecule type: DNA
A;Residues: 1-61 <FLO>
A;Cross-references: UNIPROT:P17591; GB:M23195; NID:g516584; PIDN:AAA42436.1; PID:g516584

C;Superfamily: adenovirus early E3 6.4K protein
C;Keywords: early protein

Query Match 83.3%; Score 30; DB 1; Length 61;
Best Local Similarity 80.0%; Pred. No. 60;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 20 YPHHG 24

RESULT 10
H71311
hypothetical protein TP0539 - syphilis spirochete
C;Species: Treponema pallidum subsp. pallidum (syphilis spirochete)
C;Date: 24-Jul-1998 #sequence_revision 24-Jul-1998 #text_change 09-Jul-2004
C;Accession: H71311
R;Fraser, C.M.; Norris, S.J.; Weinstock, G.M.; White, O.; Sutton, G.G.; Dodson, R.; Gwinther, J.; Khaliq, H.; Richardson, D.; Howell, J.K.; Chidambaram, M.; Utterback, T.; McDorsey, L.; Weidman, J.; Smith, H.O.; Venter, J.C.
Science 281, 375-388, 1998
A;Title: Complete genome sequence of Treponema pallidum, the syphilis spirochete.
A;Reference number: A71250; MUID:98332770; PMID:9665876
A;Accession: H71311
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-73 <COL>
A;Cross-references: UNIPROT:O83550; GB:AE001229; GB:AE000520; NID:g3322827; PIDN:AAC65533
A;Experimental source: strain Nichols
C;Genetics:
A;Gene: TP0539

Query Match 83.3%; Score 30; DB 2; Length 73;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 20 HPPH 23

RESULT 11
T47876
hypothetical protein T4C21.60 - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004
C;Accession: T47876
R;Choisne, N.; Robert, C.; Brottier, P.; Wincker, P.; Cattolico, L.; Artiguenave, F.; Sali, W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.
submitted to the Protein Sequence Database, March 2000
A;Reference number: Z24479
A;Accession: T47876
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-107 <CHO>
A;Cross-references: UNIPROT:Q9LZ28; EMBL:AL162295
A;Experimental source: cultivar Columbia; BAC clone T4C21
C;Genetics:
A;Map position: 3
A;Note: T4C21.60

Query Match 83.3%; Score 30; DB 2; Length 107;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 102 HPPH 105

RESULT 12
A03861

hypothetical protein F-112 - human adenovirus 2
C;Species: Mastadenovirus h2 (human adenovirus 2)
C;Date: 31-Mar-1992 #sequence_revision 31-Mar-1992 #text_change 09-Jul-2004
C;Accession: G92351; G92352; A03861
R;Gingeras, T.R.; Sciaky, D.; Gelinas, R.E.; Bing-Dong, J.; Yen, C.E.; Kelly, M.M.; Bull
J. Biol. Chem. 257, 13475-13491, 1982
A;Title: Nucleotide sequences from the adenovirus-2 genome.
A;Reference number: A92351; MUID:83056843; PMID:7142161
A;Accession: G92351
A;Molecule type: DNA
A;Residues: 1-112 <GIN>
A;Cross-references: UNIPROT:P03289
R;Alestrom, P.; Akuejärvi, G.; Pettersson, M.; Pettersson, U.
J. Biol. Chem. 257, 13492-13498, 1982
A;Title: DNA sequence analysis of the region encoding the terminal protein and the hypob
A;Reference number: A92352; MUID:83056844; PMID:7142162
A;Accession: G92352
A;Molecule type: DNA
A;Residues: 1-112 <ALE>

Query Match 83.3%; Score 30; DB 2; Length 112;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 27 HPPH 30

RESULT 13
B72629
hypothetical protein APE1490 - Aeropyrum pernix (strain K1)
C;Species: Aeropyrum pernix
C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C;Accession: B72629
R;Kawarayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K
DNA Res. 6, 83-101, 1999
A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A;Reference number: A72450; MUID:99310339; PMID:10382966
A;Accession: B72629
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-113 <KAW>
A;Cross-references: UNIPROT:Q9YBWO; DDBJ:AP000061; NID:95104821; PIDN:BAA80488.1; PID:d1
A;Experimental source: strain K1
C;Genetics:
A;Gene: APE1490
C;Superfamily: Aeropyrum pernix hypothetical protein APE1490

Query Match 83.3%; Score 30; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 108 HPPH 111

RESULT 14
S66936
probable membrane protein YOR053w - yeast (Saccharomyces cerevisiae)
N;Alternate names: hypothetical protein O2799
C;Species: Saccharomyces cerevisiae
C;Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 09-Jul-2004
C;Accession: S66936; S66927
R;Bohn, C.; Bolotin-Fukuhara, M.; Daignan-Fornier, B.; Dang, D.V.; Valens, M.
A;Reference number: S66929
A;Accession: S66936
A;Molecule type: DNA
A;Residues: 1-113 <BOH>
A;Cross-references: UNIPROT:Q08428; EMBL:Z74961; NID:g1420188; PID:e252336; PID:g1420189

A;Experimental source: strain S288C
R;Iandt, O.; Hiesel, R.; Unseld, M.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S66907
A;Accession: S66927
A;Molecule type: DNA
A;Residues: 1-8 <LAN>
A;Cross-references: EMBL:Z74961; MIPS:YOR053W
A;Experimental source: strain S288C
C;Genetics:
A;Cross-references: SGD:S0005579
A;Map position: 15R
C;Keywords: transmembrane protein
F;3-13/Domain: transmembrane #status predicted <TMM>

Query Match 83.3%; Score 30; DB 2; Length 113;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 90 HPPH 93

RESULT 15
TI7301
hypothetical protein DKFZp564B147.1 - human
C;Species: Homo sapiens (man)
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: TI7301
R;Bloeker, H.; Boecker, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, September 1999
A;Reference number: Z18724
A;Accession: TI7301
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-115 <BLO>
A;Cross-references: UNIPROT:Q9UFN1; EMBL:AL117556
A;Experimental source: fetal brain; clone DKFZp564B147
C;Genetics:
A;Note: DKFZp564B147.1
C;Superfamily: human hypothetical protein DKFZp564B147.1

Query Match 83.3%; Score 30; DB 2; Length 115;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
Db 55 HPPH 58

RESULT 16
MGMSB2
beta-2-microglobulin precursor - mouse
N;Alternate names: class I histocompatibility antigen beta chain
C;Species: Mus musculus (house mouse)
C;Date: 31-Mar-1981 #sequence_revision 03-Aug-1984 #text_change 09-Jul-2004
C;Accession: A90822; A93889; B93889; A12344; I79322; I58971; I63179; I63180; I63178; I631
R;Farnes, J.R.; Seidman, J.G.
Cell 29, 661-669, 1982
A;Title: Structure of wild-type and mutant mouse beta-2-microglobulin genes.
A;Reference number: A90822; MUID:83001942; PMID:6180834
A;Accession: A90822
A;Molecule type: DNA
A;Residues: 1-119 <PRN>
A;Cross-references: UNIPROT:P01887
A;Experimental source: strain BALB/c
R;Gates III, F.T.; Coligan, J.E.; Kindt, T.J.
Proc. Natl. Acad. Sci. U.S.A. 78, 554-558, 1981
A;Title: Complete amino acid sequence of murine beta-2-microglobulin: structural evidence
A;Reference number: A93889; MUID:81223779; PMID:6165997
A;Accession: A93889

A:Molecule type: protein
A:Residues: 21-119 <GAT>
A:Experimental source: strain BALB/c
A:Accession: B93889
A:Molecule type: protein
A:Residues: 21-104,'A',106-119 <GA2>
A:Experimental source: strain C57BL/6
R:Appella, E.; Tanigaki, N.; Natori, T.; Pressman, D.
Biochem. Biophys. Res. Commun. 70, 425-430, 1976
A:Title: Partial amino acid sequence of mouse beta-2-microglobulin.
A:Reference number: A12344; MUID:76231543; PMID:59595
A:Accession: A12344
A:Molecule type: protein
A:Residues: 21-48,'E',50-53,'X',55,'Z',57,'BL',60 <APP>
A:Experimental source: strain A/J
R:Parnes, J.R.; Velan, B.; Felsenfeld, A.; Ramanathan, L.; Ferrini, U.; Appella, E.; Sch
Proc. Natl. Acad. Sci. U.S.A. 78, 2253-2257, 1981
A:Title: Mouse beta-2-microglobulin cDNA clones: A screening procedure for cDNA clones o
A:Reference number: I58971; MUID:81223857; PMID:6166005
A:Accession: I79322
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 86-104,'A',106-119 <PAR>
A:Cross-references: GB:M10416; NID:g199147; PIDN:AAA39537.1; PID:g387432
A:Accession: I58971
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 31-57 <RE2>
A:Cross-references: GB:M10389; NID:g199146; PIDN:AAA39536.1; PID:g199149
R:Hermel, E.; Robinson, P.J.; Lindahl, K.F.
Immunogenetics 38, 106-116, 1993
A:Title: Sequence divergence of B2m alleles of wild Mus musculus and Mus spretus implies
A:Reference number: I49537; MUID:93246292; PMID:8482575
A:Accession: I63179
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-7,'I',9-63,'E',65-119 <PA1>
A:Cross-references: GB:M84364; NID:g199575; PIDN:AAA39668.1; PID:g199576
A:Accession: I63180
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-53,'Q',55-104,'V',106-119 <PA2>
A:Cross-references: GB:M84365; NID:g199577; PIDN:AAA39669.1; PID:g199578
A:Accession: I63178
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 24-100,'T',102-104,'V',106-115 <PA3>
A:Cross-references: GB:M84363; NID:g199573; PIDN:AAA39667.1; PID:g199574
A:Accession: I63181
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 24-53,'Q',55-104,'V',106-115 <PA4>
A:Cross-references: GB:M84366; NID:g199579; PIDN:AAA39670.1; PID:g199580
A:Accession: I63182
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-104,'V',106-119 <PA5>
A:Cross-references: GB:M84367; NID:g199581; PIDN:AAA39671.1; PID:g199582
R:Daniel, F.; Morello, D.; Le Bail, O.; Chambon, P.; Cayre, Y.; Kourilsky, P.
EMBO J. 2, 1061-1065, 1983
A:Title: Structure and expression of the mouse beta 2-microglobulin gene isolated from s
A:Reference number: I48262; MUID:84028577; PMID:6354707
A:Accession: I48262
A:Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-119 <RES>
A:Cross-references: EMBL:X01838; NID:g50104; PIDN:CAA25970.1; PID:g50105
R:Seubert, P.; Vigo-Pelfrey, C.; Esch, F.; Lee, M.; Dovey, H.; Davis, D.; Sinha, S.; Sch
Nature 359, 325-327, 1992
A:Title: Isolation and quantification of soluble Alzheimer's beta-peptide from biological
A:Reference number: S29076; MUID:93024877; PMID:1406936
A:Accession: S29077

A:Molecule type: protein
A:Residues: 'X',22-30,'X',32-44 <SEU>
A:Note: this protein was found as a minor contaminant during the purification of proteins
C:Genetics:
A:Introns: 23/1; 116/1
C:Complex: heterodimer with MHC class I histocompatibility antigen alpha chain
C:Function:
A:Description: necessary for the expression of MHC class I histocompatibility antigen
C:Superfamily: immunoglobulin C region; immunoglobulin homology
C:Keywords: heterodimer
F:1-20/Domain: signal sequence #status predicted <SIG>
F:21-119/Product: beta-2-microglobulin #status experimental <MAT>
F:38-102/Domain: immunoglobulin homology <IMM>
F:45-100/Disulfide bonds: #status predicted
Query Match 83.3%; Score 30; DB 1; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
DB 51 HPPH 54
RESULT 17
B72546
hypothetical protein APE1654 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C:Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: B72546
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ki
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: B72546
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-121 <KAW>
A:Cross-references: UNIPROT:Q9YBE3; DDBJ:AP000062; NID:g5105244; PIDN:BAA80655.1; PID:g51
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE1654
C:Superfamily: Aeropyrum pernix hypothetical protein APE1654
Query Match 83.3%; Score 30; DB 2; Length 121;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPH 4
DB 36 HPPH 39
RESULT 18
T46369
hypothetical protein DKFZp434I0118.1 - human
C:Species: Homo sapiens (man)
C:Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46369
R:Ottenwaelder, B.; Obermaier, B.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.
submitted to the Protein Sequence Database, January 2000
A:Reference number: Z23031
A:Accession: T46369
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-122 <AAA>
A:Cross-references: UNIPROT:Q9NT54; EMBL:AL137525
A:Experimental source: adult testis; clone DKFZp434I0118
C:Genetics:
A:Note: DKFZp434I0118.1
C:Superfamily: human hypothetical protein DKFZp434I0118.1

Query Match	83.3%;	Score 30;	DB 2;	Length 122;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches	4;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 HPPH 4			
Db	14 HPPH 17			
RESULT 19				
A05160				
hypothetical protein 49 - fruit fly (<i>Drosophila melanogaster</i>)				
C:Species: <i>Drosophila melanogaster</i>				
C>Date: 05-Jun-1987 #sequence_revision 05-Jun-1987 #text_change 16-Feb-1997				
C:Accession: A05160				
R:O'Connell, P.; Roobaah, M.				
Nucleic Acids Res. 12, 5495-5513, 1984				
A:Title: Sequence, structure, and codon preference of the <i>Drosophila</i> ribosomal protein 4				
A:Reference number: A93527; MUID:84272233; PMID:6087289				
A:Accession: A05160				
A:Molecule type: DNA				
A:Residues: 1-125 <OCON>				
C:Genetics:				
A:Gene: FlyBase:M(3)99D				
A:Cross-references: FlyBase:FBgn0002626				
A:Map position: 3R, 99D				
Query Match	83.3%;	Score 30;	DB 2;	Length 125;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches	4;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 HPPH 4			
Db	32 HPPH 35			
RESULT 20				
B69320				
conserved hypothetical protein AF0562 - <i>Archaeoglobus fulgidus</i>				
C:Species: <i>Archaeoglobus fulgidus</i>				
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004				
C:Accession: B69320				
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson				
; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.				
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.				
Nature 390, 364-370, 1997				
A:Authors: Uterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.				
Smith, H.O.; Woese, C.R.; Venter, J.C.				
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae				
A:Reference number: A69250; MUID:98049343; PMID:9389475				
A:Accession: B69320				
A:Status: preliminary; nucleic acid sequence not shown; translation not shown				
A:Molecule type: DNA				
A:Residues: 1-126 <KLE>				
A:Cross-references: UNIPROT:O29691; GB:AE001066; GB:AE000782; NID:g26899389; PIDN:AAB9067				
Query Match	83.3%;	Score 30;	DB 2;	Length 126;
Best Local Similarity	100.0%;	Pred. No. 1.2e+02;		
Matches	4;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;
QY	1 HPPH 4			
Db	82 HPPH 85			
RESULT 21				
F71070				
hypothetical protein PH1258 - <i>Pyrococcus horikoshii</i>				
C:Species: <i>Pyrococcus horikoshii</i>				
C>Date: 14-Aug-1998 #sequence_revision 14-Aug-1998 #text_change 12-Jul-2004				
C:Accession: F71070				
R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Hailkawa, Y.; Hino, Y.; Yamamoto, S.; Sekin				
M.; Ohfuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi				

DNA Res. 5, 55-76, 1998
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic a
 A:Reference number: A71000; MUID:98344137; PMID:9679194
 A:Accession: F71070
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-141 <KAW>
 A:Cross-references: UNIPROT:O58995; GB:AP000005; NID:g3236132; PIDN:BA30360.1; PID:d1031
 A:Experimental source: strain OT3
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank
 C:Genetics:
 A:Gene: PH1258

Query Match 83.3%; Score 30; DB 2; Length 141;
 Best Local Similarity 100.0%; Pred. No. 1.4e+02;
 Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
 Db 65 HPPH 68

RESULT 22

S39206

proline-rich protein V-beta 1 precursor - rat

C:Species: *Rattus norvegicus* (Norway rat)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C:Accession: A53118; S39206

R:Court, Y.; Rosinski-Chupin, I.; Rougeon, F.

J. Biol. Chem. 269, 520-527, 1994

A:Title: A new proline-rich protein precursor expressed in the salivary glands of the rat

A:Reference number: A53118; MUID:94103265; PMID:8276845

A:Accession: A53118

A:Status: translated from GB/EMBL/DBDJ

A:Molecule type: mRNA

A:Residues: 1-148 <COU>

A:Cross-references: UNIPROT:Q64371; EMBL:X74229; NID:g433616; PIDN:CAAS2300.1; PID:g43361

A:Note: submitted to the EMBL Data Library, July 1993

C:Genetics:

A:Gene: VCS-beta1

A:Introns: 18/3

C:Superfamily: proline-rich peptide P-B

C:Keywords: glycoprotein

F:1-18/Domain: signal sequence #status predicted <SIG>

F:19-148/Product: proline-rich protein V-beta 1 #status predicted <MAT>

F:133,143/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 83.3%; Score 30; DB 2; Length 148;

Best Local Similarity 100.0%; Pred. No. 1.5e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4

Db 97 HPPH 100

RESULT 23

A99226

conserved hypothetical protein [imported] - *Sulfolobus solfataricus*C:Species: *Sulfolobus solfataricus*

C>Date: 24-May-2001 #sequence_revision 24-May-2001 #text_change 09-Jul-2004

C:Accession: A99226

R:She, Q.; Singh, R.K.; Confalonieri, F.; Zivanovic, Y.; Allard, G.; Awayez, M.J.; Chan-y

Jong, I.; Jeffries, A.C.; Kozera, C.J.; Medina, N.; Peng, X.; Thi-Ngoc, H.P.; Redder, P.

arrett, R.A.; Ragan, M.A.; Senses, C.W.; Van der Oost, J.

submitted to GenBank, April 2001

A:Description: *Sulfolobus solfataricus* complete genome.

A:Reference number: A99139

A:Accession: A99226

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-157 <KUP>

A:Cross-references: UNIPROT:Q9UXF4; GB:AE006641; NID:g13813940; PIDN:AAK41064.1; GSPDB:G

Best Local Similarity 100.0%; Pred. No. 1.7e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HPPH 4 		
Db	43 HPPH 46		
RESULT 29 S14257 benzodiazepine receptor, peripheral-type - human C:Species: Homo sapiens (man) C>Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 12-Jul-2004 C:Accession: S14257 R:Riond, J.; Mattei, M.G.; Kaghad, M.; Dumont, X.; Guillemot, J.C.; le Fur, G.; Caput, D. Eur. J. Biochem. 195, 305-311, 1991 A:Title: Molecular cloning and chromosomal localization of a human peripheral-type benzodiazepine receptor A:Reference number: S14257; MUID:91146565; PMID:1847678 A:Accession: S14257 A>Status: preliminary A:Molecule type: mRNA A:Residues: 1-169 <RIO> A:Cross-references: UNIPROT:P30536; GB:M36035; NID:g184333; PIDN:AAA03652.1; PID:g306883 C:Genetics: A:Gene: GDB:BZRP A:Cross-references: GDB:127347; OMIM:109610 A:Map position: 22q13.31-22qter C:Superfamily: peripheral-type benzodiazepine receptor/signal transduction protein Tspo C:Keywords: mitochondrion; transmembrane protein			
Query Match 83.3%; Score 30; DB 2; Length 169; Best Local Similarity 100.0%; Pred. No. 1.7e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HPPH 4 		
Db	43 HPPH 46		
RESULT 30 T51469 glycine/proline-rich protein - Arabidopsis thaliana N:Alternate names: protein K10A8_130 C:Species: Arabidopsis thaliana (mouse-ear cress) C>Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004 C:Accession: T51469 R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew submitted to the Protein Sequence Database, August 2000 A:Reference number: Z25394 A:Accession: T51469 A>Status: preliminary A:Molecule type: DNA A:Residues: 1-173 <SAT> A:Cross-references: UNIPROT:Q9LFS9; EMBL:AL391151 A:Experimental source: Cultivar Columbia; BAC clone K10A8 C:Genetics: A:Map position: 5 A:Introns: 97/1 A>Note: K10A8_130			
Query Match 83.3%; Score 30; DB 2; Length 173; Best Local Similarity 80.0%; Pred. No. 1.7e+02; Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HPPHG 5 :		
Db	25 YPPHG 29		
RESULT 31 T52409 ucylacyanin II [imported] - Arabidopsis thaliana C:Species: Arabidopsis thaliana (mouse-ear cress)			

C>Date: 24-Oct-2000 #sequence_revision 24-Oct-2000 #text_change 09-Jul-2004 C:Accession: T52409 R:Nersissian, A.M.; Mehrabian, Z.B.; Nalbandyan, R.M.; Hart, P.J.; Fraczkiwicz, G.; Czei Protein Sci. 5, 2184-2192, 1996 A:Title: Cloning, expression, and spectroscopic characterization of Cucumis sativus stellj A:Reference number: Z17046; MUID:97084803; PMID:8931137 A:Accession: T52409 A>Status: preliminary A:Molecule type: translated from GB/EMBL/DBDJ A:Residues: 1-181 <NER> A:Cross-references: UNIPROT:O80517; EMBL:U76299; PIDN:AAC32039.1 C:Superfamily: stellacyanin			
Query Match 83.3%; Score 30; DB 2; Length 181; Best Local Similarity 100.0%; Pred. No. 1.8e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HPPH 4 		
Db	156 HPPH 159		
RESULT 32 B97784 formylmethionine deformylase (EC.3.5.1.31) - Rickettsia conorii (strain Malish 7) C:Species: Rickettsia conorii C>Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004 C:Accession: B97784 R:Ogata, H.; Audic, S.; Renesto-Audiffren, P.; Fournier, P.E.; Barbe, V.; Samson, D.; Ro Science 293, 2093-2098, 2001 A:Title: Mechanisms of Evolution in Rickettsia conorii and Rickettsia prowazekii. A:Reference number: A97700; MUID:21442074; PMID:11557893 A:Accession: B97784 A>Status: preliminary A:Molecule type: DNA A:Residues: 1-183 <KUR> A:Cross-references: UNIPROT:Q92HU7; GB:AE006914; PIDN:AAL03212.1; PID:g15619763; GSPDB:GB C:Genetics: A:Gene: def2 C:Superfamily: peptidase deformylase C:Keywords: hydrolase			
Query Match 83.3%; Score 30; DB 2; Length 183; Best Local Similarity 100.0%; Pred. No. 1.8e+02; Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 HPPH 4 		
Db	172 HPPH 175		
RESULT 33 AC2903 conserved hypothetical protein Atu2660 [imported] - Agrobacterium tumefaciens (strain C58 C:Species: Agrobacterium tumefaciens C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004 C:Accession: AC2903 R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, L. erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell ; Karp, P.; Romero, P.; Zhang, S. Science 294, 2317-2323, 2001 A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, I ster, E.W. A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58. A:Reference number: AB2577; MUID:21608550; PMID:11743193 A:Accession: AC2903 A>Status: preliminary A:Molecule type: DNA A:Residues: 1-183 <KUR> A:Cross-references: UNIPROT:O8UC38; GB:AE008688; PIDN:AAL43641.1; PID:g17741163; GSPDB:GB A:Experimental source: strain C58 (Dupont) C:Genetics: A:Gene: Atu2660			

A;Map position: circular chromosome

Query Match 83.3%; Score 30; DB 2; Length 183;
Best Local Similarity 100.0%; Pred. No. 1.8e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
DB 43 HPPH 46

RESULT 34

H69937

hypothetical protein yjpj - Bacillus subtilis

C;Species: Bacillus subtilis
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C;Accession: H69937

R;Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berteaux-Lecq, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.; Koester, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois, A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauseel, Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yagumoto, K.; Yata, K.; Yoshida, K.; Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A. Bacillus subtilis.
A;Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A;Reference number: A69580; PMID:98044033; PMID:9384377

A;Accession: H69937
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-203 <KUN>

A;Cross-references: UNIPROT:P54172; GB:T299115; GB:AL009126; NID:G2634478; PIDN:CAB14102.
A;Experimental source: strain 168

C;Genetics:

A;Gene: yjpj

C;Superfamily: Bacillus subtilis hypothetical protein yjpj

Query Match 83.3%; Score 30; DB 2; Length 203;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
DB 162 HPPH 165

RESULT 35

B84315

hypothetical protein Vngl625h [imported] - Halobacterium sp. NRC-1

C;Species: Halobacterium sp. NRC-1

C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C;Accession: B84315

R;Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berquist, B.; Pan, M.; Shukla, H.D.; Laaky, S.; Leithausen, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablon, J.; Jung, K.H.; Alam, M.; Freitas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A;Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; Li
A;Title: Genome sequence of Halobacterium species NRC-1.

A;Reference number: B84160; PMID:20504483; PMID:11016950

A;Accession: B84315

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-207 <STO>

A;Cross-references: UNIPROT:Q9HPT1; GB:AE004437; NID:g10581098; PIDN:AAG19886.1; GSPDB:C

C;Genetics:

A;Gene: VNG1625H

Query Match 83.3%; Score 30; DB 2; Length 207;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
DB 23 HPPH 26

RESULT 36

D75004

hypothetical protein PAB1037 - Pyrococcus abyssi (strain Orsay)

N;Alternate names: yafJ homolog

C;Species: Pyrococcus abyssi

C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C;Accession: D75004

R;Anonymous, Genoscope

submitted to the EMBL Data Library, July 1999

A;Description: Pyrococcus abyssi genome sequence: insights into archaeal chromosome struc

A;Reference number: A75001

A;Accession: D75004

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-208 <KAW>

A;Cross-references: UNIPROT:Q9UYD8; GB:AJ248288; GB:AL056836; NID:g5458960; PIDN:CAB50474

A;Experimental source: strain Orsay

C;Genetics:

A;Gene: PAB1037

Query Match 83.3%; Score 30; DB 2; Length 208;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
DB 198 HPPH 201

RESULT 37

C89005

protein T24A6.3 [imported] - Caenorhabditis elegans

C;Species: Caenorhabditis elegans

C;Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001

C;Accession: C89005

R;Anonymous, The C. elegans Sequencing Consortium.

Science 282, 2012-2018, 1998

A;Title: Genome sequence of the nematode C. elegans: a platform for investigating biology

A;Reference number: A75000; PMID:99069613; PMID:9851916

A;Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg

A;Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and

A;Accession: C89005

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-209 <STO>

A;Cross-references: GB:chr_v; PIDN:AAC17789.1; PID:g3168937; GSPDB:GN00023; CESP:T24A6.3

C;Genetics:

A;Gene: T24A6.3

A;Map position: 5

Query Match 83.3%; Score 30; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4
|||||
DB 148 HPPH 151

RESULT 38

AZ2901

conserved hypothetical protein Atu2647 [imported] - Agrobacterium tumefaciens (strain C5)

C;Species: Agrobacterium tumefaciens

C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004

C;Accession: AF2901
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AF2901
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-214 <KUR>
A;Cross-references: UNIPROT:Q8UC51; GB:AE008688; PIDN:AAL43628.1; PID:gl7741149; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu2647
A;Map position: circular chromosome

Query Match 83.3%; Score 30; DB 2; Length 214;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4

||||

Db 17 HPPH 20

RESULT 39

H97676

hypothetical protein AGR_C.4799 [imported] - Agrobacterium tumefaciens (strain C58, Cere

C;Species: Agrobacterium tumefaciens

C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004

C;Accession: H97676

R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quorollo, B.; Goldman,

A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz, B.;

Science 294, 2323-2328, 2001

A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum

A;Reference number: A97359; MUID:21608551; PMID:11743194

A;Accession: H97676

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-214 <KUR>

A;Cross-references: UNIPROT:Q8UC51; GB:AE007869; PIDN:AAK88369.1; PID:gl5157854; GSPDB:G

C;Genetics:

A;Gene: AGR_C.4799

A;Map position: circular chromosome

Query Match

83.3%; Score 30; DB 2; Length 214;

Best Local Similarity 100.0%; Pred. No. 2.1e+02;

Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4

||||

Db 17 HPPH 20

RESULT 40

I53100

eHAND - mouse

C;Species: Mus sp. (mouse)

C;Date: 02-Aug-1996 #sequence_revision 02-Aug-1996 #text_change 05-Nov-1999

C;Accession: I53100

R;Cserrjesi, P.; Brown, D.; Lyons, G.E.; Olson, E.N.

Dev. Biol. 170, 664-678, 1995

A;Title: Expression of the novel basic helix-loop-helix gene eHAND in neural crest deriv

A;Reference number: I53100; MUID:95377552; PMID:7649392

A;Accession: I53100

A;Status: preliminary; translated from GB/EMBL/DBJ

A;Molecule type: mRNA

A;Residues: 1-216 <RES>

A;Cross-references: GB:S79216; NID:gl086931; PIDN:AAB35104.1; PID:gl086932

Query Match 83.3%; Score 30; DB 2; Length 216;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPH 4

||||

Db 17 HPPH 20

Search completed: June 15, 2005, 14:22:49

Job time : 19.5 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-9

Perfect score: 36

Sequence: 1 HPPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	36	100.0	62	Q15952	Q15952 homo sapien
2	36	100.0	108	Q655P3	Q655P3 oryza sativ
3	36	100.0	118	Q9YDU9	Q9YDU9 aeropyrum p
4	36	100.0	171	Q7PG70	Q7PG70 anopheles g
5	36	100.0	176	Q9KMP0	Q9KMP0 vibrio chol
6	36	100.0	193	Q8FJG7	Q8FJG7 escherichia
7	36	100.0	195	Q52719	Q52719 rhodobacter
8	36	100.0	214	Q8PQU3	Q8PQU3 xanthomonas
9	36	100.0	234	Q7USZ6	Q7USZ6 rhodospirell
10	36	100.0	235	PRPX HORVU	P16273 hordeum vul
11	36	100.0	240	Q9XH33	Q9XH33 oryza sativ
12	36	100.0	240	Q9XH36	Q9XH36 oryza sativ
13	36	100.0	251	Q9Y9K5	Q9Y9K5 aeropyrum p
14	36	100.0	277	P96427	P96427 rhodococcus
15	36	100.0	311	Q9YUN5	Q9YUN5 equid herpe
16	36	100.0	321	Q8FT73	Q8FT73 corynebacte
17	36	100.0	355	Q73RX1	Q73RX1 mycobacteri
18	36	100.0	399	Q6MTV5	Q6MTV5 aspergillus
19	36	100.0	449	Q89615	Q89615 clostridium
20	36	100.0	470	1 RHBA RHIME	Q923r2 r diamnoba
21	36	100.0	491	Q73TJ7	Q73TJ7 mycobacteri
22	36	100.0	493	Q6P5Y4	Q6P5Y4 homo sapien
23	36	100.0	503	Q9W259	Q9W259 drosophila
24	36	100.0	526	1 HRG RABIT	Q28640 oryctolagus
25	36	100.0	560	2 Q9YHE8	Q9YHE8 brachydanio
26	36	100.0	581	2 Q87RZ6	Q87RZ6 vibrio para
27	36	100.0	583	2 Q8DB43	Q8DB43 vibrio vuln
28	36	100.0	587	2 Q7MMD1	Q7MMD1 vibrio vuln
29	36	100.0	601	2 Q75HV8	Q75HV8 oryza sativ
30	36	100.0	723	2 Q86BB8	Q86BB8 drosophila
31	36	100.0	736	2 Q6C276	Q6C276 yarrowia li

32	36	100.0	773	2	Q95R83	Q95R83 drosophila
33	36	100.0	938	2	Q86BB9	Q86BB9 drosophila
34	36	100.0	1075	2	Q6XLI4	Q6XLI4 ciona intes
35	36	100.0	1427	2	Q91562	Q91562 xenopus lae
36	36	100.0	1445	2	Q63155	Q63155 rattus norv
37	36	100.0	1447	1	DCC HUMAN	P43146 homo sapien
38	36	100.0	1447	1	DCC MOUSE	P70211 mus musculu
39	36	100.0	2013	2	Q65ZE5	Q65ZE5 cochllobolu
40	30	83.3	16	2	Q9S8A0	Q9S8A0 pinus monti
41	30	83.3	30	2	Q9V2V2	Q9V2V2 methanobact
42	30	83.3	30	2	Q7LYR7	Q7LYR7 plasmid pme
43	30	83.3	54	2	Q8KM82	Q8KM82 mycoplasma
44	30	83.3	55	2	Q9NDZ7	Q9NDZ7 leishmania
45	30	83.3	61	1	E306_ADE35	P17591 human adeno

ALIGNMENTS

RESULT 1

Q15952 ID Q15952 PRELIMINARY; PRT; 62 AA.
AC Q15952;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-NOV-1996 (Tremblrel. 01, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Agrin (fragment).
GN Name=AGRN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92407628; PubMed=1326608;
RA Rupp F., Ocelik T., Linial M., Peterson K., Francke U., Scheller R.;
RT "Structure and chromosomal localization of the mammalian agrin gene.";
RL J. Neurosci. 12:3535-3544(1992).
DR EMBL; S44195; AAB23327.1; -.
DR HSSP; P21674; 1LR7.
DR GO; GO:0005605; C:basal lamina; TAS.
DR GO; GO:0007268; P:synaptic transmission; NAS.
DR InterPro; IPR003645; FOL N.
DR InterPro; IPR002350; Prot_inh_Kazal.
DR Pfam; PF00050; Kazal_1; 1.
DR SMART; SM00274; FOLN; 1.
DR SMART; SM00280; KAZAL; 1.
FT NON_TER 62 62
SQ SEQUENCE 62 AA; 6344 MW; 4CD1653B6B8F09E8 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 62;
Best Local Similarity 100.0%; Pred. No. 33;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 26 HPPHG 30

RESULT 2

Q655P3 ID Q655P3 PRELIMINARY; PRT; 108 AA.
AC Q655P3;
DT 25-OCT-2004 (Tremblrel. 28, Created)
DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)
DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)
DE Hypothetical protein B1011A07.27.
GN Name=B1011A07.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;

```

RN SEQUENCE FROM N.A.
RX PubMed=1247438;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mirubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AF003722; BAD45462.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11886 MW; 6B93CD836B85C0A7 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 29 HPPHG 33

RESULT 3
Q9YDU9 PRELIMINARY; PRT; 118 AA.
ID Q9YDU9
AC Q9YDU9
DT 01-NOV-1999 (TREMBLrel. 12, Created)
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)
DE Hypothetical protein APE0819.
GN OrderedLocusNames=APE0819;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA79798.1; -.
DR FIR; F72674; F72674.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 118 AA; 12953 MW; 6134441D6406CDB0 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 118;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 83 HPPHG 87

RESULT 4

```

```

Q7PG70 PRELIMINARY; PRT; 171 AA.
ID Q7PG70
AC Q7PG70;
DT 01-MAR-2004 (TREMBLrel. 26, Created)
DT 01-MAR-2004 (TREMBLrel. 26, Last sequence update)
DE ENSANGP0000022414 (Fragment).
GN Name=ENSANGG0000020585;
OS Anopheles gambiae str. PEST.
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Anopheles.
OX NCBI_TaxID=180454;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PEST;
RA Anopheles Genome Sequencing Consortium;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
CC -! CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AAB0100847; EAA45048.1; -.
FT NON_TER 1 171
FT SEQUENCE 171 AA; 18450 MW; 1A135D70C5F14B67 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 171;
Best Local Similarity 100.0%; Pred. No. 90;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 111 HPPHG 115

RESULT 5
Q9KMP0 PRELIMINARY; PRT; 176 AA.
ID Q9KMP0
AC Q9KMP0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE Hypothetical protein VCA0284.
GN OrderedLocusNames=VCA0284;
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=El Tor N16961 / Serotype O1;
RX MEDLINE=20406833; PubMed=10952301; DOI=10.1038/35020000;
RA Heidelberg J.P., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Unayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Rettelink H., Richardson D.L.,
RA Ermlaeva M.D., Vamathevan J.J., Bass S., Qin H., Dragoi I.,
RA Sellers P., McDonald L.A., Utterback T.R., Fleischmann R.D.,
RA Nierman W.C., White O., Salzberg S.L., Smith H.O., Colwell R.R.,
RA Mekalanos J.J., Venter J.C., Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae.";
RL Nature 406:477-483(2000).
DR EMBL; AE004368; AAF96193.1; -.
DR FIR; F82478; F82478.
DR TIGR; VCA0284; -.
DR InterPro; IPR008727; PAAR.
DR Pfam; PF05488; PAAR_motif; 2.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 176 AA; 18431 MW; 79A633E68C447515 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 HPPHG 5
Db      51 HPPHG 55

RESULT 6
Q8FJG7
ID      Q8FJG7      PRELIMINARY;      PRT;      193 AA.
AC      Q8FJG7;
DT      01-MAR-2003 (TReMBLrel. 23, Created)
DT      01-MAR-2003 (TReMBLrel. 23, Last sequence update)
DE      Hypothetical protein c0967.
GN      OrderedLocuNames=c0967;
OS      Escherichia coli O6.
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC      Enterobacteriaceae; Escherichia.
OX      NCBI_TaxID=217992;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=06:HI / CFT073 / ATCC 700928 / UPEC;
RX      MEDLINE=22388234; PubMed=12471157; DOI=10.1073/pnas.252529799;
RA      Wasko R.A., Burland V., Plunkett G. III, Redford P., Roesch P.,
RA      Ravelo D., Buckles E.L., Liou S.-R., Boutin A., Hackett J., Stroud D.,
RA      Mayhew G.F., Rose D.J., Zhou S., Schwartz D.C., Perna N.T.,
RA      Mobley H.L.T., Donnenberg M.S., Blattner F.R.;
RT      "Extensive mosaic structure revealed by the complete genome sequence
RT      of uropathogenic Escherichia coli.";
RL      Proc. Natl. Acad. Sci. U.S.A. 99:17020-17024 (2002).
DR      EMBL; AE016758; AAN79440.1; -.
KW      Complete proteome.
SQ      SEQUENCE 193 AA; 21398 MW; B3EAECS743D88D3D CRC64;

Query Match      100.0%; Score 36; DB 2; Length 193;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      117 HPPHG 121

RESULT 7
Q52719
ID      Q52719      PRELIMINARY;      PRT;      195 AA.
AC      Q52719;
DT      01-NOV-1996 (TReMBLrel. 01, Created)
DT      01-NOV-1996 (TReMBLrel. 01, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      Orf10 (Fragment).
OS      Rhodobacter capsulatus (Rhodospseudomonas capsulata).
OC      Bacteria; Proteobacteria; Alphaproteobacteria; Rhodobacterales;
OC      Rhodobacteraceae; Rhodobacter.
OX      NCBI_TaxID=1061;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=B10S;
RX      MEDLINE=90384835; PubMed=2402451;
RA      Jouanneau Y., Richaud P., Grabau C.;
RT      "The nucleotide sequence of a flavodoxin-like gene which precedes two
RT      ferredoxin genes in Rhodobacter capsulatus.";
RL      Nucleic Acids Res. 18:5284-5284 (1990).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=B10S;
RX      MEDLINE=90036712; PubMed=2681157;
RA      Schatt E., Jouanneau Y., Vignais P.M.;
RT      "Molecular cloning and sequence analysis of the structural gene of
RT      ferredoxin I from the photosynthetic bacterium Rhodobacter
RT      capsulatus.";
RL      J. Bacteriol. 171:6218-6226 (1989).
RN      [3]
RP      SEQUENCE FROM N.A.

```

```

RC      STRAIN=B10S;
RX      MEDLINE=91302301; PubMed=2071578;
RA      Saeki K., Suetesugu Y., Tokuda K.-i., Miyatake Y., Young D.A.,
RA      Marrs B.L., Matsubara H.;
RT      "Genetic analysis of functional differences among distinct ferredoxins
RT      in Rhodobacter capsulatus.";
RL      J. Biol. Chem. 266:12889-12895 (1991).
RN      [4]
RP      SEQUENCE FROM N.A.
RC      STRAIN=B10S;
RX      MEDLINE=94088454; PubMed=8264535;
RA      Schmehl M., Jahn A., Meyer zu Vilsendorf A., Hennecke S., Masepohl B.,
RA      Schuppeler M., Marxer M., Oelze J., Klipp W.;
RT      "Identification of a new class of nitrogen fixation genes in
RT      Rhodobacter capsulatus: a putative membrane complex involved in
RT      electron transport to nitrogenase.";
RL      Mol. Gen. Genet. 241:602-615 (1993).
DR      EMBL; X72888; CAA51406.1; -.
FT      NON TER 195
SQ      SEQUENCE 195 AA; 20985 MW; 56D1AFA4ADEE7D7C CRC64;

Query Match      100.0%; Score 36; DB 2; Length 195;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 HPPHG 5
Db      116 HPPHG 120

RESULT 8
Q8PQU3
ID      Q8PQU3      PRELIMINARY;      PRT;      214 AA.
AC      Q8PQU3;
DT      01-OCT-2002 (TReMBLrel. 22, Created)
DT      01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT      01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE      Hypothetical protein XAC0228.
GN      OrderedLocuNames=XAC0228;
OS      Xanthomonas axonopodis (pv. citri).
OC      Bacteria; Proteobacteria; Gammaproteobacteria; Xanthomonadales;
OC      Xanthomonadaceae; Xanthomonas.
OX      NCBI_TaxID=92829;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=306 / ATCC 13902 / XV 101;
RX      MEDLINE=22024215; PubMed=12024217; DOI=10.1038/417459a;
RA      da Silva A.C.R., Ferro J.A., Reinach F.C., Farah C.S., Furlan L.R.,
RA      Quaggio R.B., Monteiro-Vitorello C.B., Van Sluys M.A., Almeida N.F.,
RA      Alves L.M.C., do Amaral A.M., Bertolini M.C., Camargo L.E.A.,
RA      Camarotte G., Cannavan F., Cardozo J., Chambergo F., Ciapina L.P.,
RA      Cicarelli R.M.B., Coutinho L.L., Cursino-Santos J.R., El-Dorry H.,
RA      Faria J.B., Ferreira A.J.S., Ferreira R.C.C., Ferro M.I.T.,
RA      Formighieri E.F., Franco M.C., Greggio C.C., Gruber A.,
RA      Katsuyama A.M., Kishi L.T., Leite R.P., Lemos E.G.M., Lemos M.V.F.,
RA      Locati E.C., Machado M.A., Madeira A.M.B.N., Martinez-Rossi N.M.,
RA      Martins E.C., Meidanis J., Menck C.F.M., Miyaki C.Y., Moon D.H.,
RA      Moreira L.M., Novo M.T.M., Okura V.K., Oliveira M.C., Oliveira V.R.,
RA      Pereira H.A., Roasi A., Sena J.A.D., Silva C., de Souza R.F.,
RA      Spinola L.A.F., Takita M.A., Tamura R.E., Teixeira E.C., Tezza R.I.D.,
RA      Trindade dos Santos M., Truffi D., Tsai S.M., White F.F.,
RA      Setubal J.C., Kitajima J.P.;
RT      "Comparison of the genomes of two Xanthomonas pathogens with differing
RT      host specificities.";
RL      Nature 417:459-463 (2002).
DR      EMBL; AE011647; AAM35120.1; -.
KW      Complete proteome.
SQ      SEQUENCE 214 AA; 22782 MW; 439DB42F4C1258E4 CRC64;

Query Match      100.0%; Score 36; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```
QY 1 HPPHG 5
DB 199 HPPHG 203

RESULT 9
Q7US26 PRELIMINARY; PRT; 234 AA.
AC Q7US26;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS OrderedLocusNames=RB4213;
GN Rhodopirellula baltica.
OC Bacteria; Planctomycetes; Planctomycetacia; Planctomycetales;
OC Planctomycetaceae; Pirellula.
OX NCBI_TaxID=117;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1;
RX MEDLINE=22735913; PubMed=12835416; DOI=10.1073/pnas.1431443100;
RA Gloeckner F.O., Kube M., Bauer M., Teeling H., Lombardot T.,
RA Ludwig W., Gade D., Beck A., Borzym K., Heitmann K., Rabus R.,
RA Schlesner H., Amann R., Reinhardt R.;
RT "Complete genome sequence of the marine planctomycete Pirellula sp.
RT strain 1.";
RL Proc. Natl. Acad. Sci. U.S.A. 100:8298-8303 (2003).
DR EMBL; BX294140; CAD73646.1; --
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 234 AA; 26924 MW; 03A974166BAA3FEC CRC64;

Query Match 100.0%; Score 36; DB 2; Length 234;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 135 HPPHG 139

RESULT 10
PRPX_HORVU STANDARD; PRT; 235 AA.
AC P16273;
DT 01-AUG-1990 (Rel. 15, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Pathogen-related protein.
OS Hordeum vulgare (Barley).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Triticeae; Hordeum.
OX NCBI_TaxID=4513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Psakon 4; TISSUE=Leaf;
RX MEDLINE=90067951; PubMed=258721;
RA Jutidamrongphan W., Mackinnon G., Manners J.M., Scott K.J.;
RT "Sequence of a near-full length cDNA clone for a mRNA of barley
RT induced by fungal infection.";
RL Nucleic Acids Res. 17:9478-9478 (1989).
RN [2]
RP REVISIONS.
RC STRAIN=cv. Psakon 4; TISSUE=Leaf;
RA Scott K.J.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
CC -!- INDUCTION: By fungus Bipolaris sorokiniana infection.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
```

```
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; X16648; CAA34641.1; --
DR PIR; S09604; S09604.
DR PIR; T06168; T06168.
KW Pathogenesis-related protein; Plant defense.
SQ SEQUENCE 235 AA; 26293 MW; 995136DD24A9CD3C CRC64;

Query Match 100.0%; Score 36; DB 1; Length 235;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 176 HPPHG 180

RESULT 11
Q9XH3 PRELIMINARY; PRT; 240 AA.
AC Q9XH3;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative pathogenesis related protein.
GN Name=OSJNBa0049B20.20; Synonyms=P0034C11.23;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R., Benito M.-I., Lin X., Mason T.M., Umayam L., Shea T.P.,
RA Fujii C.Y., Shen M., Fraser C.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316 (2002).
DR EMBL; AC007789; AAD38293.1; --
DR EMBL; AP02865; BAB18332.1; --
DR Gramene; Q9XH3; --
SQ SEQUENCE 240 AA; 27196 MW; 4ABC7B300FAFEBD6 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 177 HPPHG 181

RESULT 12
Q9XH6
```

ID Q9XH6 PRELIMINARY; PRT; 240 AA.
AC Q9XH6;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Putative pathogenesis related protein.
GN Names:OSJNBa0049820.17; Synonyms:P0034C11.20;
OS Oryza sativa (japonica cultivar-group)
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzoideae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell R., Benito M.-I., Lin X., Mason T.M., Umeyam L., Shea T.P.,
RA Fujii C.Y., Shen M., Fraser C.M.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilmura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Itonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AC007789; RAD38290.1; -;
DR EMBL; AF002865; BAB18330.1; -;
DR Genome; Q9XH6; -;
SQ SEQUENCE 240 AA; 27146 MW; 86C7D7A6F409AD7E CRC64;

Query Match 100.0%; Score 36; DB 2; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 178 HPPHG 182

RESULT 13
Q9Y9K5 PRELIMINARY; PRT; 251 AA.
AC Q9Y9K5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein APE2283.
GN OrderedLocNames=APE2283;
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=KI;
RX MEDLINE=99310339; PubMed=10382986;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;

RT "Complete genome sequence of an aerobic hyper-thermophilic
crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000064; BAA81295.1; -;
DR FIR; G72454; G72454.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR004843; M-pesterase.
DR Pfam; PF00149; Metallophos; 1.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 251 AA; 27075 MW; 9C30F5292F51B60 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 251;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 172 HPPHG 176

RESULT 14
P96427 PRELIMINARY; PRT; 277 AA.
AC P96427;
DT 01-MAY-1997 (TREMBlrel. 03, Created)
DT 01-MAY-1997 (TREMBlrel. 03, Last sequence update)
DT 01-MAY-2004 (TREMBlrel. 26, Last annotation update)
DE AraC-type regulatory protein.
GN Name=fasR;
OS Rhodococcus fascians.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Nocardiaceae; Rhodococcus.
OX NCBI_TaxID=1828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D188;
RX MEDLINE=20461232; PubMed=11004184;
RX DOI=10.1128/JB.182.20.5832-5840.2000;
RA Temmerman W., Vereecke D., Dreesen R., Van Montagu M., Holsters M.,
RA Goethals K.;
RT "Leafy gall formation is controlled by fasR, an AraC-type regulatory
gene in Rhodococcus fascians.";
RL J. Bacteriol. 182:5832-5840(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=D188;
RA Vereecke D.M.;
RL Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Y09820; CAA70950.1; -;
DR GO; GO:0005622; C:intracellular; IEA.
DR GO; GO:0003700; P:transcription factor activity; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009057; Homeodomain_like.
DR InterPro; IPR000005; HTHARAC.
DR Pfam; PF00165; HTH_Arac; 2.
DR PRINTS; PR00032; HTHARAC.
DR SMART; SM00342; HTH_ARAC; 1.
DR PROSITE; PS01124; HTH_ARAC_FAMILY_2; 1.
KW DNA-binding; Transcription; Transcription regulation.
SQ SEQUENCE 277 AA; 30866 MW; 071DB9DEFCL1E113 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 277;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
Db 155 HPPHG 159

RESULT 15
Q9YJN5 PRELIMINARY; PRT; 311 AA.
ID Q9YJN5

AC O9YJN5;
 DT 01-MAY-1999 (TrEMBLrel. 10, Created)
 DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein (CARD-like apoptotic protein) (V-CARMEN).
 GN Name=CLAP; Synonyms=E10;
 OS Equid herpesvirus 2.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Gammaherpesvirinae; Rhadinovirus.
 OX NCBI_TaxID=12657;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99142601; PubMed=989495; DOI=10.1016/S0092-8674(00)80957-5;
 RX Williams T.G., Jadavel D.M., Du M.Q., Peng H., Perry A.R.,
 RA Abdul-Rauf M., Price H., Karan L., Majekodunmi O., Wlodarska I.,
 RA Pan L., Crook T., Hamoudi R., Isaacson P., Dyer M.J.;
 RT "Bcl10 is involved in t(1;14)(p22;q32) of MALT B-cell lymphoma and
 RT mutated in multiple tumor types.";
 RL Cell 96:35-45(1999).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP Williams A.G.;
 RA Submitted (MAY-1998) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99292766; PubMed=10364242; DOI=10.1074/jbc.274.25.17946;
 RX Srinivasula S.M., Ahmad M., Lin J.H., Poyet J.L.,
 RA Fernandes-Alnemri T., Tsichlis P.N., Alnemri E.S.;
 RT "CLAP, a novel caspase recruitment domain-containing protein in the
 RT tumor necrosis factor receptor pathway, regulates NF-kappaB activation
 RT and apoptosis.";
 RL J. Biol. Chem. 274:17946-17954(1999).
 RL [4]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=99214546; PubMed=10187771; DOI=10.1074/jbc.274.15.9962;
 RX Thome M., Martinon F., Hofmann K., Rubio V., Steiner V., Schneider P.,
 RA Mattmann C., Tschopp J.;
 RT "Equine herpesvirus-2 E10 gene product, but not its cellular
 RT homologue, activates NF-kappaB transcription factor and c-Jun N-
 RT terminal kinase.";
 RL J. Biol. Chem. 274:9962-9968(1999).
 DR EMBL; AJ006410; CAA07016.1; -.
 DR EMBL; AF134394; AAD39146.1; -.
 DR EMBL; AF100340; AAD16430.1; -.
 DR GO; GO:0005622; C:intracellular; IEA.
 DR GO; GO:0005515; F:protein binding; IEA.
 DR GO; GO:0042981; P:regulation of apoptosis; IEA.
 DR InterPro; IPR001315; CARD.
 DR InterPro; IPR011029; DEATH_like.
 DR Pfam; PF00619; CARD; 1.
 DR PROSITE; PS0209; CARD; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 311 AA; 32602 MW; 04EC68C7353D723B CRC64;
 Query Match 100.0%; Score 36; DB 2; Length 311;
 Best Local Similarity 100.0%; Pred. No. 1.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 DB 218 HPPHG 222
 RESULT 16
 Q8FT73 PRELIMINARY; PRT; 321 AA.
 AC Q8FT73;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Putative oxppcycle protein OpcA.
 GN OrderedLocusNames=CE1697;
 OS Corynebacterium efficiens.

OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Corynebacteriaceae; Corynebacterium.
 OX NCBI_TaxID=152794;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=YS-314;
 RX MEDLINE=22723752; PubMed=12840036; DOI=10.1101/gr.1285603;
 RA Nishio Y., Nakamura Y., Kawarabayashi Y., Uguda Y., Kimura E.,
 RA Sugimoto S., Matsui K., Yamagishi A., Kikuchi H., Ikeo K.,
 RA Gojobori T.;
 RT "Comparative complete genome sequence analysis of the amino acid
 RT replacements responsible for the thermostability of Corynebacterium
 RT efficiens.";
 RL Genome Res. 13:1572-1579(2003).
 DR EMBL; AP005219; BAC18507.1; -.
 KW Complete proteome.
 SQ SEQUENCE 321 AA; 34772 MW; 2278243DEDD90920 CRC64;
 Query Match 100.0%; Score 36; DB 2; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 DB 187 HPPHG 191
 RESULT 17
 Q73RX1 PRELIMINARY; PRT; 355 AA.
 AC Q73RX1;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=MAP4307;
 OS Mycobacterium paratuberculosis.
 OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
 OC Corynebacterineae; Mycobacteriaceae; Mycobacterium.
 OX NCBI_TaxID=1770;
 RN [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=k10;
 RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
 RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE017242; AAS06857.1; -.
 DR InterPro; IPR001865; Ribosomal_S2.
 DR PROSITE; PS00962; RIBOSOMAL_S2_1; UNKNOWN_1.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 355 AA; 38394 MW; A97DD1B43A5797D CRC64;
 Query Match 100.0%; Score 36; DB 2; Length 355;
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 HPPHG 5
 DB 286 HPPHG 290
 RESULT 18
 Q6MYV5 PRELIMINARY; PRT; 399 AA.
 AC Q6MYV5;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Possible bhlh transcription factor.
 GN ORFNames=Afa28D10.070;
 OS Aspergillus fumigatus (Sartorya fumigata).
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;
 OC Eurotiales; Trichocomaceae; mitosporic trichocomaceae; Aspergillus.
 OX NCBI_TaxID=5085;

```

RN  RP  SEQUENCE FROM N.A.
RX  Pain A., Woodward J., Quail M.A., Anderson M.J., Clark R., Collins M.,
RA  Fosker N., Fraser A., Harris D., Larkie N., Murphy L., Humphray S.,
RA  O'Neill S., Perlea M., Price C., Rabinowitch E., Rajandream M-A.,
RA  Salzberg S., Saunders D., Seagar K., Sharp S., Warren T.,
RA  Denning D.W., Barrell B., Hall N.,
RT  "Insight into the genome of Aspergillus fumigatus: analysis of a 922
RT  kb region encompassing the nitrate assimilation gene cluster."
RL  Fungal Genet. Biol. 41:443-453(2004).
CC  -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.
DR  EMBL; BX649605; CAB47904.1; -.
DR  InterPro; IPR001092; HLH_basic.
DR  Pfam; PF00010; HLH; 1.
DR  SMART; SM00353; HLH; 1.
DR  PROSITE; PS00898; HLH; 1.
SQ  SEQUENCE 399 AA; 42514 MW; D23C914E17F6F15C CRC64;

Query Match 100.0%; Score 36; DB 2; Length 399;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 310 HPPHG 314

RESULT 19
ID Q89615 PRELIMINARY; PRT; 449 AA.
AC Q89615;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE RnfC/nqrF.
GN Name=rnfC; OrderedLocusNames=CTC01019;
OS Clostridium tetani.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1513;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baumeier S., Fricke W.F., Wieser A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Massachusetts / E88;
RX MEDLINE=22457253; PubMed=12552129; DOI=10.1073/pnas.0335853100;
RA Brueggemann H., Baumeier S., Fricke W.F., Wieser A., Liesegang H.,
RA Decker I., Herzberg C., Martinez-Arias R., Merkl R., Henne A.,
RA Gottschalk G.;
RT "The genome sequence of Clostridium tetani, the causative agent of
RT tetanus disease."
RL Proc. Natl. Acad. Sci. U.S.A. 100:1316-1321(2003).
DR EMBL; AB015939; AAC35605.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005489; F:electron transporter activity; IEA.
DR GO; GO:0005506; F:iron ion binding; IEA.
DR GO; GO:0006118; P:electron transport; IEA.
DR InterPro; IPR001450; 4FeS_ferredoxin.
DR InterPro; IPR011538; Complex1 51k dom.
DR InterPro; IPR009051; Helical_ferredox.
DR InterPro; IPR010208; RnfC.
DR InterPro; IPR011054; Rndmt_hyb_motif.
DR Pfam; PF01512; Complex1_51k; 1.
DR Pfam; PF00037; Fer4; 1.
DR PRINTS; PR00353; 4FEASFRDOXIN.
DR TIGRFAMs; TIGR01945; rnfC; 1.
DR PROSITE; PS00198; 4FEAS_FERREDOXIN; 2.
KW 4Fe-4S; Complete proteome; iron; iron-sulfur; Metal-binding.
SQ SEQUENCE 449 AA; 49052 MW; 6F98DC0B98F52C78 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 449;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 25 HPPHG 29

RESULT 20
RHBA RHIME
ID RHBA RHIME STANDARD; PRT; 470 AA.
AC Q923R2;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Diaminobutyrate-2-oxoglutarate aminotransferase (EC 2.6.1.76) (L-
DE diaminobutyric acid transaminase) (DABA-AT) (L-2,4-diaminobutyrate:2-ketoglutarate 4-
DE aminotransferase) (DABA-AT)
DE Name=rhBA; Synonyms=rhSA; OrderedLocusNames=RA1258; ORFNames=SMA2400;
GN Rhizobium meliloti (Sinorhizobium meliloti).
OS Plasmid pSYMA (megaplasmid 1).
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RCR2011 / SU47;
RX MEDLINE=21172875; PubMed=11274118;
RX DOI=10.1128/JB.183.8.2576-2585.2001;
RA Lynch D., O'Brien J., Welch T., Clarke P., Cuiv P.O., Crosa J.H.,
RA O'Connell M.;
RT "Genetic organization of the region encoding regulation, biosynthesis,
RT and transport of rhizobactin 1021, a siderophore produced by
RT Sinorhizobium meliloti."
RL J. Bacteriol. 183:2576-2585 (2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubier F., Bowser L., Capela D., Galibert F., Gouzy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSYMA megaplasmid."
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
CC -1- CATALYTIC ACTIVITY: L-2,4-diaminobutyrate + 2-oxoglutarate = L-
CC glutamate + L-aspartic 4-semialdehyde.
CC -1- COFACTOR: Pyridoxal phosphate.
CC -1- PATHWAY: Siderophore biosynthesis; rhizobactin biosynthesis.
CC -1- SIMILARITY: Belongs to the class-III pyridoxal-phosphate-dependent
CC aminotransferase family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF110737; AAD09412.1; -.
CC EMBL; AB007311; AAK65916.1; -.
CC PIR; B95419; B95419.
CC PIR; T46814; T46814.
CC HSP; P12995; 1QJ3.
CC InterPro; IPR005814; Aminotrans_3.
CC InterPro; IPR004637; Dat.
CC Pfam; PF00202; Aminotran_3; 1.
CC TIGRFAMs; TIGR00709; dat; 1.
CC TIGRFAMs; TIGR01885; Orn aminotrans; 1.
CC PROSITE; PS00500; AA_TRANSF_CLASS_3; 1.
KW Aminotransferase; Complete proteome; Iron transport; Plasmid;
KW Pyridoxal phosphate; Transferrase.

```

FT BINDING 304 304 Pyridoxal phosphate (Potential).
SQ SEQUENCE 470 AA; 50148 MW; 0377B38B9A09049 CRC64;

Query Match 100.0%; Score 36; DB 1; Length 470;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 406 HPPHG 410

RESULT 21

ID Q73TJ7 PRELIMINARY; PRT; 491 AA.
AC Q73TJ7;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=MAP3721;
OS Mycobacterium paratuberculosis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1770;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K10;
RA Li L., Bannantine J., Zhang Q., Amonsin A., Alt D., Kapur V.;
RL Submitted (SEP-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB017240; AA06271.1; -;
KW Complete proteome.
SQ SEQUENCE 491 AA; 54266 MW; 442E24F9A031D8E0 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 11 HPPHG 15

RESULT 22

ID Q6P5Y4 PRELIMINARY; PRT; 493 AA.
AC Q6P5Y4;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
RA Brownstein M.J., Uedin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McSwain P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Morley D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalhus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Lung;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC062581; AA062581.1; -;
DR InterPro; IPR003599; IG.
DR InterPro; IPR007110; IG-like.
DR InterPro; IPR003597; IG cl.
DR InterPro; IPR003598; IG cl.
DR InterPro; IPR003596; IG v.
DR Pfam; PF00047; IG 2;
DR SMART; SM00409; IG; 4.
DR SMART; SM00407; IG cl; 2.
DR SMART; SM00408; IG cl; 4.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PS50835; IG LIKE; 4.
KW Hypothetical protein.
SQ SEQUENCE 493 AA; 52761 MW; CD7C5591CC4822D2 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 493;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
|||||
Db 477 HPPHG 481

RESULT 23

Q9W259 PRELIMINARY; PRT; 503 AA.
ID Q9W259;
AC Q9W259;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG13508-PA.
GN ORFNames=CG13506;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Aghayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,
RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris K.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,


```
FT SITE 421 422 Cleavage (by plasmin).
SQ SEQUENCE 526 AA; 58877 MW; 810F23D367D93D42 CRC64;
Query Match 100.0%; Score 36; DB 1; Length 526;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 353 HPPHG 357
RESULT 25
QYHE8 PRELIMINARY; PRT; 560 AA.
AC Q9YHE8;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE HMG-box transcription factor (Transcription factor 7-like 1a) (T-cell
DE specific, HMG-box).
GN Name:tcf7l1a; Synonyms=TCF-3, tcf3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Fellegri F., Maischein H.-M.;
RT "Function of zebrafish beta-catenin and TCF-3 in dorsoventral
RT patterning.";
RL Mech. Dev. 77:0-0(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalek A.U., Smailus D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Embryo;
RA Strausberg R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF097410; AAC64609.1; -
DR EMBL; BC053135; AAH53135.1; -
DR HSP; P27782; 2LEF
DR ZFIN; ZDB-GENE-980605-30; tcf7l1a.
DR GO; GO:0003677; F:DNA binding; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009071; HMG-box.
DR Pfam; PF00505; HMG_box; 1.
DR SMART; SM00398; HMG; 1.
DR PROSITE; PS01118; HMG_BOX_2; 1.
SQ SEQUENCE 560 AA; 61459 MW; DC35C50EA4D61232 CRC64;
Query Match 100.0%; Score 36; DB 2; Length 560;
Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 278 HPPHG 282
RESULT 26
Q87RZ6 PRELIMINARY; PRT; 581 AA.
AC Q87RZ6;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Hypothetical protein VP0628.
GN OrderedLocustNames=VP0628;
OS Vibrio parahaemolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=670;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=RIMD 2210633 / Serotype O3:K6;
RX MEDLINE=22508454; PubMed=12620739; DOI=10.1016/S0140-6736(03)12659-1;
RA Makino K., Oshima K., Kurokawa K., Yokoyama K., Uda T., Tagomori K.,
RA Iijima Y., Nijima M., Nakano M., Yamashita A., Kubota Y., Kimura S.,
RA Yasunaga T., Honda T., Shinagawa H., Hattori M., Iida T.;
RT "Genome sequence of Vibrio parahaemolyticus: a pathogenic mechanism
RT distinct from that of V. cholerae.";
RL Lancet 361:743-749(2003).
DR EMBL; AP005075; BAC58891.1; -
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_like.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_like; 1.
DR Complete proteome; Hypothetical protein.
RW SEQUENCE 581 AA; 64511 MW; 292D0C34F2159F9E CRC64;
Query Match 100.0%; Score 36; DB 2; Length 581;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 HPPHG 5
Db 348 HPPHG 352
RESULT 27
Q8D843 PRELIMINARY; PRT; 583 AA.
AC Q8D843;
DT 01-MAR-2003 (TREMBlrel. 23, Created)
DT 01-MAR-2003 (TREMBlrel. 23, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Predicted metal-dependent hydrolase with the TIM-barrel fold.
GN OrderedLocustNames=VV13145;
OS Vibrio vulnificus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=672;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CMCP6;
RA Rhee J.H., Kim S.Y., Chung S.S., Kim J.J., Moon Y.H., Jeong H.,
RA Choy H.E.;
RT "Complete genome sequence of Vibrio vulnificus CMCP6.";
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
```

```

DR EMBL; AE016807; AA01465.1; -.
DR GO: GO:0016787; F.hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_1.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_1; 1.
DR Complete proteome; Hydrolase.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 583 AA; 64278 MW; 64AF798F8DCE69A6 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 583;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 348 HPPHG 352

RESULT 28
ID Q7MMD1 PRELIMINARY; PRT; 587 AA.
AC Q7MMD1;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Predicted metal-dependent hydrolase.
GN OrderedLocusNames=VVI1141;
OS Vibrio vulnificus (strain YJ016).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=196600;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14656965; DOI=10.1101/gr.1295503;
RA Chen C.-Y., Wu K.-M., Chang Y.-C., Chang C.-H., Tsai H.-C.,
RA Liao T.-L., Liu Y.-M., Chen H.-J., Shen A.B.-T., Li J.-C., Su T.-L.,
RA Shao C.-P., Lee C.-T., Hor L.-I., Tsai S.-F.;
RT "Comparative genome analysis of Vibrio vulnificus, a marine
RT pathogen.";
RL Genome Res. 13:2577-2587(2003).
DR EMBL; AP005334; BAC93905.1; -.
DR GO: GO:0016787; F.hydrolase activity; IEA.
DR InterPro; IPR006680; Amidohydro_1.
DR InterPro; IPR011550; Amidohydro_1.
DR InterPro; IPR011059; Metallo_hydrolase.
DR Pfam; PF01979; Amidohydro_1; 1.
DR ProDom; PD001248; Amidohydro_1; 1.
DR Complete proteome; Hydrolase.
KW Complete proteome; Hydrolase.
SQ SEQUENCE 587 AA; 64736 MW; B94C7FA5B3D15E7F CRC64;

Query Match 100.0%; Score 36; DB 2; Length 587;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 352 HPPHG 356

RESULT 29
ID Q7SHV8 PRELIMINARY; PRT; 601 AA.
AC Q7SHV8;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypochemical protein OSUNB0092621.7;
GN Name=OSUNB0092621.7;
OS Oryza sativa (Japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzae; Oryza.

```

```

OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Chow T.-Y., Hsing Y.-I.C., Chen C.-S., Chen H.-H., Liu S.-M.,
RA Chao Y.-T., Chang S.-J., Chen H.-C., Chen S.-K., Chen T.-R.,
RA Chen Y.-L., Cheng C.-H., Chung C.-I., Han S.-Y., Hsiao S.-H.,
RA Hsiung J.-N., Hsu C.-H., Huang J.-J., Kau P.-I., Lee M.-C.,
RA Li Y.-F., Lin S.-J., Lin Y.-C., Wu S.-W., Yu C.-Y., Yu S.-W.,
RA Wu H.-P., Shaw J.-F.;
RL Submitted (MAY-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC134932; AA07640.1; -.
KW Hypothetical protein.
SQ SEQUENCE 601 AA; 66009 MW; B78E8DAC1AFFFD2B CRC64;

Query Match 100.0%; Score 36; DB 2; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 5 HPPHG 9

RESULT 30
ID Q86BB8 PRELIMINARY; PRT; 723 AA.
AC Q86BB8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG30492-PD.
GN ORFNames=CG30492;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA De Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svikas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weisenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,

```



```

Q86BB9 ID Q86BB9 PRELIMINARY; PRT; 938 AA.
AC Q86BB9;
DT 01-JUN-2003 (TReMBLrel. 24, Created)
DT 01-JUN-2003 (TReMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE CG30492-PE.
GN ORFNAMES=CG30492;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwe R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bortova D., Borchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtova C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glöck A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wasmann D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodgett, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426065; PubMed=12537568;
RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
RA George R.A., Hoskins R.A., Laverty T., Muzny D.M., Nelson C.R.,
RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
RA Svirskas R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
RA Weinstein G., Scher S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
RT melanogaster euchromatic genome sequence."
RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426070; PubMed=12537573;
RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
RA Ashburner M., Celniker S.E.;
RT "The transposable elements of the Drosophila melanogaster euchromatin:

```

```

RT a genomics perspective."
RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
RN [4]
RP SEQUENCE FROM N.A.
RX MEDLINE=22426069; PubMed=12537572;
RA Mira S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
RA Lewis S.E.;
RT "Annotation of the Drosophila melanogaster euchromatic genome: a
RT systematic review."
RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
RN [5]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RG FlyBase;
RL Submitted (MAR-2004) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AE003840; AAO41496.1; -.
DR FlyBase; FBgn0050492; CG30492.
DR InterPro; IPR011011; FYVE_PHD_ZnF.
DR InterPro; IPR011009; Kinase like.
SQ SEQUENCE 938 AA; 102954 MW; 9984B6D579A1EA14 CRC64;
Query Match 100.0%; Score 36; DB 2; Length 938;
Best Local Similarity 100.0%; Pred. No. 4.8e+02; Mismatches 0; Indels 0; Gaps 0;
Matches 5; Conservative 0;

QY 1 HPPHG 5
DB 729 HPPHG 733
|||||

RESULT 34
Q6XLI4 PRELIMINARY; PRT; 1075 AA.
AC Q6XLI4;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TReMBLrel. 27, Last annotation update)
DE Swan.
OS Ciona intestinalis.
OC Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
OC Phlebobranchia; Clonidae; Ciona.
OX NCBI_TaxID=7719;
RN [1]
RP SEQUENCE FROM N.A.
RA Chen Y., Huang C.-H.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AY226102; AAP48572.1; -.
DR InterPro; IPR000504; RNA_rec_mot.
DR Pfam; PF00076; RRM 1; 3.
DR SMART; SM00360; RRM; 6.
DR PROSITE; PS0102; RRM; 3.
SQ SEQUENCE 1075 AA; 120486 MW; 24A993DB6F48CBD8 CRC64;
Query Match 100.0%; Score 36; DB 2; Length 1075;
Best Local Similarity 100.0%; Pred. No. 5.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 880 HPPHG 884
|||||

RESULT 35
Q91562 PRELIMINARY; PRT; 1427 AA.
ID Q91562

```

```
AC Q91562;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Tumor suppressor.
GN Name=XDCCa;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95113183; PubMed=7813784; DOI=10.1006/dbio.1994.1345;
RA Pierceall W.E., Reale M.A., Candia A.F., Wright C.V., Cho K.R.,
RA Fearon E.R.;
RT "Expression of a homologue of the deleted in colorectal cancer (DCC)
RT gene in the nervous system of developing Xenopus embryos.";
RL Dev. Biol. 166:654-665(1994).
DR EMBL; U10986; AAA70168.1; -.
DR PIR; I51669; I51669.
DR HSSP; Q9UQH9; 1EV2.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; ig; 1.
DR PRINTS; PR00014; FNTYPEI11.
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS50835; IG_LIKE; 4.
DR SEQUENCE 1427 AA; 156533 MW; 61FEA12C8A674972 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 1427;
Best Local Similarity 100.0%; Pred. No. 7.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 1066 HPPHG 1070

RESULT 36
Q63155 PRELIMINARY; PRT; 1445 AA.
AC Q63155;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DE Deleted in Colorectal Cancer.
GN Name=DCC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97015074; PubMed=8861902; DOI=10.1016/S0092-8674(00)81336-7;
RA Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.-Y.,
RA Culotti J.G., Tessier-Lavigne M.;
RA Oliner J.D., Hedge P., Silverman G.A., Vogelstein B.;
RT "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RL Cell 87:175-185(1996).
DR EMBL; U68725; AAB41099.1; -.
DR HSSP; P13596; IQZ1.
DR InterPro; IPR003962; FNIII subd.
DR InterPro; IPR003961; FN_III.
DR InterPro; IPR008957; FN_III-like.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003598; Ig_c2.
DR InterPro; IPR010560; Neogenin_C.
DR Pfam; PF00041; fn3; 6.
DR Pfam; PF00047; ig; 1.
DR Pfam; PF06583; Neogenin_C; 1.
DR PRINTS; PR00014; FNTYPEI11.
```

```
DR SMART; SM00060; FN3; 6.
DR SMART; SM00408; IGC2; 3.
DR PROSITE; PS50853; FN3; 6.
DR PROSITE; PS50835; IG_LIKE; 4.
SQ SEQUENCE 1445 AA; 157940 MW; 084F625954481988 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 1445;
Best Local Similarity 100.0%; Pred. No. 7.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
DB 1084 HPPHG 1088

RESULT 37
DCC_HUMAN STANDARD; PRT; 1447 AA.
AC P43146;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Netrin receptor DCC precursor (tumor suppressor protein DCC)
DE (Colorectal cancer suppressor).
GN Name=DCC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95011532; PubMed=7926722;
RA Hedrick L., Cho K.R., Fearon E.R., Wu T.-C., Kinzler K.W.,
RA Vogelstein B.;
RT "The DCC gene product in cellular differentiation and colorectal
RT tumorigenesis.";
RL Genes Dev. 8:1174-1183(1994).
RN [2]
RP SEQUENCE OF 1-750 FROM N.A.
RX MEDLINE=90100559; PubMed=2294591;
RA Fearon E.R., Cho K.R., Nigro J.M., Kern S.E., Simons J.W.,
RA Ruppert J.M., Hamilton S.R., Preisinger A.C., Thomas G., Kinzler K.W.,
RA Vogelstein B.;
RT "Identification of a chromosome 18q gene that is altered in colorectal
RT cancers.";
RL Science 247:49-56(1990).
RN [3]
RP SEQUENCE OF 107-472 FROM N.A. (SCRAMBELD EXONS).
RX MEDLINE=91121517; PubMed=1991322; DOI=10.1016/0092-8674(91)90244-S;
RA Nigro J.M., Cho K.R., Fearon E.R., Kern S.E., Ruppert J.M.,
RA Oliner J.D., Kinzler K.W., Vogelstein B.;
RT "Scrambled exons.";
RL Cell 64:607-613(1991).
RN [4]
RP GENE STRUCTURE, AND VARIANT CARCINOMA H1S-1375.
RX MEDLINE=94245241; PubMed=8188295;
RA Cho K.R., Oliner J.D., Simons J.W., Hedrick L., Fearon E.R.,
RA Preisinger A.C., Hedge P., Silverman G.A., Vogelstein B.;
RT "The DCC gene: structural analysis and mutations in colorectal
RT carcinomas.";
RL Genomics 19:525-531(1994).
RN [5]
RP FUNCTION.
RX MEDLINE=97015074; PubMed=8861902; DOI=10.1016/S0092-8674(00)81336-7;
RA Keino-Masu K., Masu M., Hinck L., Leonardo E.D., Chan S.S.,
RA Culotti J.G., Tessier-Lavigne M.;
RA Oliner J.G., Tessier-Lavigne M.;
RT "Deleted in Colorectal Cancer (DCC) encodes a netrin receptor.";
RL Cell 87:175-185(1996).
RN [6]
RP INTERACTIONS WITH SIAH1 AND SIAH2, AND DEGRADATION.
RX PubMed=9334332;
RA Hu G., Zhang S., Vidal M., Baer J.L., Xu T., Fearon E.R.;
RT "Mammalian homologs of seven in absentia regulate DCC via the
```


axon attraction of neuronal growth cones in the developing nervous system upon ligand binding. Its association with UNC5 proteins may trigger signaling for axon repulsion. It also acts as a dependence receptor required for apoptosis induction when not associated with netrin ligand. Implicated as a tumor suppressor gene.

-!- SUBUNIT: Interacts with the cytoplasmic part of UNC5A, UNC5B, UNC5C and probably UNC5D (By similarity).

-!- SUBCELLULAR LOCATION: Type I membrane protein.

-!- ALTERNATIVE PRODUCTS:

Event=Alternative splicing; Named isoforms=2;

Name=A;

isoId=P70211-1; Sequence=Displayed;

Note=Isoform B is produced by alternative initiation at Met-85 of isoform A;

Name=C;

isoId=P70211-2; Sequence=VSP_002501;

Event=Alternative initiation;

Comment=2 isoforms, A (shown here) and B, are produced by alternative initiation at Met-1 and Met-85;

-!- TISSUE SPECIFICITY: In the embryo, expressed at high levels in the developing brain and neural tube. In adult, highly expressed in brain with very low levels found in testis, heart and thymus.

Isoform C is expressed only in the embryo.

-!- DEVELOPMENTAL STAGE: Low levels in early gestation. Highest levels expressed during mid gestation. Levels decrease in late gestation and remain at this level in the adult.

-!- PTM: Ubiquitinated; mediated by SIAH1 or SIAH2 and leading to its subsequent proteasomal degradation (By similarity).

-!- SIMILARITY: Belongs to the immunoglobulin superfamily. DCC family.

-!- SIMILARITY: Contains 6 fibronectin type III domains.

-!- SIMILARITY: Contains 4 immunoglobulin-like C2-type domains.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@sib-sib.ch).

DR EMBL; X85788; CAAS9786.1; --.

DR HSSP; P13590; LIES.

DR MGD; MGI:94869; Dcc.

DR InterPro; IPR003961; FN III.

DR InterPro; IPR008957; FN III-like.

DR InterPro; IPR003962; FNIII subd.

DR InterPro; IPR007110; Ig-like.

DR InterPro; IPR003598; Ig_c2.

DR InterPro; IPR010560; Neogenin_C.

DR Pfam; PF00041; fn3; 6.

DR Pfam; PF00047; ig; 4.

DR Pfam; PF06583; Neogenin_C; 1.

DR PRINTS; PR00014; ENTPEVII.

DR SMART; SM00060; FN3; 6.

DR SMART; SM00408; IGc2; 3.

DR PROSITE; PS00853; FN3; 6.

DR PROSITE; PS00835; IG LIKE; 4.

KW Alternative initiation; Alternative splicing; Anti-oncogene;

KW Apoptosis; Developmental protein; Glycoprotein; Immunoglobulin domain;

KW Receptor; Repeat; Signal; Transmembrane; Ubl conjugation.

FT SIGNAL 1 25 Potential.

FT CHAIN 26 1447 Netrin receptor DCC, isoform A.

FT CHAIN 85 1447 Netrin receptor DCC, isoform B.

FT INIT MET 85 85 For isoform B.

FT DOMAIN 26 1097 Extracellular (Potential).

FT TRANSMEM 1098 1122 Potential.

FT DOMAIN 1123 1447 Cytoplasmic (Potential).

FT DOMAIN 139 229 Ig-like C2-type 1.

FT DOMAIN 234 326 Ig-like C2-type 2.

FT DOMAIN 331 416 Ig-like C2-type 3.

FT DOMAIN 429 520 Ig-like C2-type 4.

FT DOMAIN 429 520 Fibronectin type-III 1.

FT DOMAIN 528 616 Fibronectin type-III 2.

FT	DOMAIN	622	714	Fibronectin type-III 3.
FT	DOMAIN	726	814	Fibronectin type-III 4.
FT	DOMAIN	843	939	Fibronectin type-III 5.
FT	DOMAIN	944	1041	Fibronectin type-III 6.
FT	DISULFID	61	117	By similarity.
FT	DISULFID	161	212	By similarity.
FT	DISULFID	261	310	By similarity.
FT	DISULFID	352	400	By similarity.
FT	CARBOHYD	60	60	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	94	94	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	299	299	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	318	318	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	478	478	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	628	628	N-linked (GlcNAc. .) (Potential).
FT	CARBOHYD	702	702	N-linked (GlcNAc. .) (Potential).
FT	VARSPPLIC	819	838	Missing (in isoform C).
FT				/FTId=VSP_002501.
SQ	SEQUENCE	1447 AA;	159298 MW;	ODIF1097C22D5B9F CRC64;

Query Match 100.0%; Score 36; DB 1; Length 1447;

Best Local Similarity 100.0%; Pred. No. 7.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1 HPPHG 5
Db	1084 HPPHG 1088

|||||

RESULT 39

Q6SLES PRELIMINARY; PRT; 2013 AA.

AC Q6SLES;

DT 05-JUL-2004 (TRENBLrel. 27, Created)

DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)

DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)

DE Putative histidine Kinase HKZp.

GN Name=HKZ;

OS Cochliobolus heterostrophus (Drechslera maydis).

OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Dothideomycetes;

OC Pleosporales; Pleosporaceae; Cochliobolus.

OX NCBI_TaxID=5016;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C4;

RX PubMed:14665450; DOI=10.1128/EC.2.6.1151-1161.2003;

RA Catlett N.L., Yoder O.C., Turgeon B.G.;

RT "Whole-genome analysis of two-component signal transduction genes in fungal pathogens.";

RL Eukaryotic Cell 2:1151-1161 (2003).

CC -!- SIMILARITY: Contains 1 histidine kinase domain.

DR EMBL; AY456005; AAR29881.1; --.

DR GO; GO:0016020; C:membrane; IEA.

DR GO; GO:0005524; F:ATP binding; IEA.

DR	GO; GO:0003677; F:DNA binding; IEA.
DR	GO; GO:0016301; F:kinase activity; IEA.
DR	GO; GO:0000156; F:two-component response regulator activity; IEA.
DR	GO; GO:0000155; F:two-component sensor molecule activity; IEA.
DR	GO; GO:0007600; P:sensory perception; IEA.
DR	GO; GO:0000160; P:two-component signal transduction system (p. .); IEA.
DR	InterPro; IPR003594; ATPbind_ATPase.
DR	InterPro; IPR004358; Bact_sens_pr_C.
DR	InterPro; IPR011006; Chey_like.
DR	InterPro; IPR005467; His Kinase.
DR	InterPro; IPR003661; His_kinA_N.
DR	InterPro; IPR001610; PAC_
DR	InterPro; IPR000014; PAS.
DR	InterPro; IPR000700; PAS-assoc_C.
DR	Pfam; PF02518; HATPase_c; 1.
DR	Pfam; PF00512; HisKA; 1.
DR	Pfam; PF00785; PAC; 2.
DR	Pfam; PF00989; PAS; 2.
DR	Pfam; PF00072; Response_reg; 1.

DR PRINTS; PRO0344; BCTRLSENSOR.
 DR ProDom; PD000039; Response; reg; 1.
 DR SMART; SM00387; HATPase_C; 1.
 DR SMART; SM00388; HsKA; 1.
 DR SMART; SM00086; PAC; 2.
 DR SMART; SM00091; PAC; 2.
 DR SMART; SM00448; REC; 1.
 DR TIGREAME; TIGR00229; sensory_box; 1.
 DR PROSITE; PS50109; HIS_KIN; 1.
 DR PROSITE; PS50113; PAC; 2.
 DR PROSITE; PS50112; PAS; 2.
 DR PROSITE; PS50110; RESPONSE REGULATORY; 1.
 KW Kinase; Phosphorylation; Sensory transduction.
 SQ SEQUENCE 2013 AA; 21997 MW; E4B99CFCC9D57065 CRC64;

Query Match 100.0%; Score 36; DB 2; Length 2013;
 Best Local Similarity 100.0%; Pred. No. 1e+03;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 DB 94 HPPHG 98

RESULT 40
 QS8A0
 ID QS8A0 PRELIMINARY; PRT; 16 AA.
 AC QS8A0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
 DE Ribulose BIPHOSPHATE carboxylase small subunit (fragment).
 OS Pinus monticola (Western white pine).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus.
 OX NCBI_TaxID=3345;
 RN [1]
 RP SEQUENCE.
 RX MEDLINE=96213005; PubMed=8665095;
 RA Ekramoddoullah A.K., Taylor D.W.;
 RT "Seasonal variation of western white pine (Pinus monticola D. Don)
 RL foliage proteins";
 RL Plant Cell Physiol. 37:189-199(1996).
 SQ SEQUENCE 16 AA; 1803 MW; 27D4934C881717DE CRC64;

Query Match 83.3%; Score 30; DB 2; Length 16;
 Best Local Similarity 80.0%; Pred. No. 1e+02;
 Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 HPPHG 5
 DB 4 HPPYG 8

Search completed: June 15, 2005, 14:21:27
 Job time : 86.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:52:01 ; Search time 90 Seconds
(without alignments)
21.487 Million cell updates/sec

Title: US-10-074-225A-10

Perfect score: 35
Sequence: 1 pPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	5	5	Abb79811 Histidine
2	35	100.0	5	8	Adh10416 Rabbit HP
3	35	100.0	17	3	Aab39278 Human sec
4	35	100.0	53	4	Abb70178 Drosophil
5	35	100.0	55	4	Aam15229 Peptide #
6	35	100.0	55	4	Abb34219 Peptide #
7	35	100.0	55	4	Aam27690 Peptide #
8	35	100.0	55	4	Abb29058 Peptide #
9	35	100.0	55	4	Abb19655 Protein #
10	35	100.0	55	4	Aam67398 Human bon
11	35	100.0	55	4	Aam55013 Human bra
12	35	100.0	55	4	Abb49057 Human liv
13	35	100.0	55	4	Aam02972 Peptide #
14	35	100.0	55	5	Abb37025 Human pep
15	35	100.0	57	4	Aau54664 Propionib
16	35	100.0	57	6	Abb51183 Propionib
17	35	100.0	60	4	Aau43160 Propionib
18	35	100.0	60	6	Abb39679 Propionib
19	35	100.0	61	4	Aam37888 Peptide #
20	35	100.0	61	4	Aam77682 Human bon
21	35	100.0	61	4	Aam64953 Human bra
22	35	100.0	61	4	Abb59329 Human liv
23	35	100.0	75	4	Aau32380 Novel hum
24	35	100.0	77	4	Aau41116 Propionib
25	35	100.0	77	4	Aau43934 Propionib

26	35	100.0	77	6	ABM40453	Abm40453 Propionib
27	35	100.0	77	6	ABM37635	Abm37635 Propionib
28	35	100.0	79	2	AAy12827	Aay12827 Human S'
29	35	100.0	80	4	ABB61183	Abb61183 Drosophil
30	35	100.0	89	5	ABP11024	Abp11024 Human ORF
31	35	100.0	89	6	ABU34317	Abu34317 Protein e
32	35	100.0	97	4	ABG24494	Abg24494 Novel hum
33	35	100.0	98	7	ABO63515	AbO63515 Klebsiell
34	35	100.0	101	8	ADH10412	Adh10412 Rabbit HP
35	35	100.0	103	7	ABO66179	AbO66179 Klebsiell
36	35	100.0	111	4	AAG67890	Aag67890 Human gly
37	35	100.0	111	4	AAG64349	Aag64349 Glycosyl
38	35	100.0	115	3	AAG16968	Aag16968 Arabidops
39	35	100.0	116	3	AAG16967	Aag16967 Arabidops
40	35	100.0	118	4	AAO01973	Aao01973 Human pol
41	35	100.0	124	4	AAO04253	Aao04253 Human pol
42	35	100.0	124	7	ADC32845	Adc32845 Human nov
43	35	100.0	124	7	ADF60519	Adf60519 Human con
44	35	100.0	129	4	AAU44880	Aau44880 Propionib
45	35	100.0	129	6	ABM41399	Abm41399 Propionib

ALIGNMENTS

RESULT 1

ABB79811	
ID	ABB79811 standard; peptide; 5 AA.
AC	ABB79811;
XX	
DT	25-NOV-2002 (first entry)
XX	
DE	Histidine proline rich glycoprotein pentapeptide.
XX	
KW	Histidine proline rich glycoprotein; HPRG; antiangiogenic; cytostatic;
KW	antiarteriosclerotic; antiinflammatory; antidiabetic; gynaecological;
KW	antiarthritic; antiulcer; osteopathic; antitumour; ophthalmological;
KW	nootropic; neuroprotective; antiparkinsonian.
OS	Synthetic.
XX	
PN	WO200264621-A2.
XX	
PD	22-AUG-2002.
XX	
PF	14-FEB-2002; 2002WO-US004336.
XX	
PR	14-FEB-2001; 2001US-0268370P.
XX	
PA	(ATTE-) ATTENUON LLC.
XX	
PI	Donate F, Harris S, Plunkett ML, Mazar AP;
XX	
DR	WPI; 2002-666989/71.
XX	
PT	New histidine-proline rich glycoprotein (HPRG) polypeptide, useful for
PT	diagnosing or treating diseases associated with undesired cell migration,
PT	invasion, proliferation, or angiogenesis, e.g. cancer, atherosclerosis.
XX	
PS	Claim 2; Page 67; 82pp; English.
XX	
CC	The present sequence is a specific example of claimed anti-angiogenic
CC	pentapeptides of the invention. Claimed anti-angiogenic polypeptides or
CC	peptides comprise: the histidine-proline-rich (H/P) domain of human
CC	histidine proline rich glycoprotein (HPRG, see ABB79806); an H/P domain
CC	inhibiting angiogenesis, endothelial cell proliferation or endothelial
CC	tube formation in vitro or in vivo; or a pentapeptide having the generic
CC	sequence given in ABB79808, such as the present peptide, or its variant
CC	having an additional 1 to 4 amino acids comprising His, Pro or Gly at its
CC	N- or C-terminus. Also claimed are: chemically synthesised or
CC	recombinantly produced peptide multimers; a diagnostically or

therapeutically labeled anti-angiogenic polypeptide, peptide or peptide multimer; a diagnostically useful HPRG-related composition, comprising the diagnostically labeled polypeptide, peptide or peptide multimer and a carrier; an antibody specific for an epitope of HPRG that is present in the H/P domain of human or rabbit HPRG, and which binds to HPRG or to any of the H/P domains in a way which inhibits the anti-angiogenic activity of HPRG or the domain, or an antigen-binding fragment of the antibody; a method for inhibiting cell migration, cell invasion, cell proliferation or angiogenesis, or for inducing apoptosis; a method for treating a subject having a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis; a method of stimulating or inhibiting angiogenesis in a subject; a method of detecting the presence of HPRG or its cleavage product or its peptide in a biological sample; isolated nucleic acids encoding the polypeptide, peptide or peptide multimer; an expression vector; transfected or transfecting cells; a method of providing to a cell, tissue or organ an angiogenesis-inhibitory amount of HPRG, an H/P domain of HPRG or its pentapeptide, or the peptide multimer; an affinity ligand useful for binding to, or isolating, an HPRG-binding molecule or cells expressing the binding molecule, comprising the polypeptide, peptide or peptide multimer, immobilised to a solid support or carrier; and a method of isolating HPRG-binding molecule, or isolating or enriching cells expressing HPRG-binding site or receptor, from a complex mixture. The compositions and methods are useful in diagnosing or treating a disease or condition associated with undesired cell migration, invasion, proliferation, or angiogenesis, such as cancer, atherosclerosis, diabetic retinopathy, inflammation, endometriosis, arthritis, peptic ulcers, or fractures. HPRG is especially useful in inhibiting the growth of primary tumours or metastases, and may also be used in treating neurodegenerative diseases like Alzheimer's or Parkinson's disease. The antibodies are stimulators of angiogenesis and are useful for promoting neovascularization in pertinent disease states, and in various immunoassays

XX Sequence 5 AA;

Query Match 100.0%; Score 35; DB 5; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPHG 5
Db 1 PPHG 5

RESULT 2
ADH10416
XX ADH10416 standard; peptide; 5 AA.

XX ADH10416;

XX 11-MAR-2004 (first entry)

XX Rabbit HPRG protein H/P rich domain repeat fragment.

XX Tpm; tropomyosin; antiangiogenic receptor;
XX histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
XX ophthalmological; antiinflammatory; gynaecological; antiarthritic;
XX antipsoriatic; dermatological; cardiant; vasotropic; vulnary;
XX angiogenesis; gene therapy; rabbit.

XX Oryctolagus cuniculus.

XX WO2003077872-A2.

XX 25-SEP-2003.

XX 17-MAR-2003; 2003WO-US008060.

XX 15-MAR-2002; 2002US-0364047P.

XX (ATTE-) ATTENUON LLC.

PI Mcrae K, Donate F, Juarez J, Mazar AP;
XX WPI; 2004-090604/09.

XX New tropomyosin-related antiangiogenic receptor polypeptide, useful for
PT inhibiting endothelial cell migration, invasion, proliferation or
PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
XX cancer.

PS Example 5; SEQ ID NO 30; 117pp; English.

XX The invention relates to an isolated tropomyosin (Tpm)-related
CC antiangiogenic receptor polypeptide or peptide, which is a fragment
CC of a full-length native Tpm protein expressed on the surface of
CC endothelial cells, or a variant of the fragment. It has a molecular mass
CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
CC an internal fragment of a native Tpm isoform which is a binding site for
CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
CC polypeptide, peptide or variant has substantially the same biochemical
CC activity of binding to the antiangiogenic polypeptide agents, as does the
CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
CC binds to the isolated polypeptide or peptide is human histidine-proline
CC rich glycoprotein (HPRG); rabbit HPRG, a Tpm-binding, antiangiogenic
CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
CC human kinogen (HK), the D5 domain of HK, or a Tpm-binding,
CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
CC antibodies and compositions are useful for inhibiting endothelial cell
CC migration, invasion, proliferation or angiogenesis, for inducing
CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
CC psoriasis, or scleroderma. The antibody may be also used for detecting
CC the presence of a Tpm polypeptide or peptide in a biological sample, for
CC promoting wound healing, or for treating diseases or conditions in which
CC increased angiogenesis is desired, e.g. coronary artery disease or
CC peripheral artery disease. The present sequence represents a repeat
CC fragment present in the rabbit HPRG protein His-Pro (H/P) rich domain.

XX Sequence 5 AA;

Query Match 100.0%; Score 35; DB 8; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PPHG 5
Db 1 PPHG 5

RESULT 3
AAB39278
ID AAB39278 standard; protein; 17 AA.

XX AAB39278;

XX 02-FEB-2001 (first entry)

XX Human secreted protein sequence encoded by gene 31 SEQ ID NO:159.

XX Human; secreted protein; immunosuppressive; antiarthritic; antirheumatic;
KW antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
KW neotropic; neuroprotective; antibacterial; virucide; fungicide; neoplasm;
KW ophthalmological; autoimmune disease; rheumatoid arthritis; angiogenesis;
KW hyperproliferative disorder; cardiovascular disorder; infection;
KW cerebrovascular disorder; nervous system disorder; ocular disorder;
KW wound healing; chemotaxis.

XX Homo sapiens.

XX WO200056754-A1.

XX 28-SEP-2000.

XX

PF 16-MAR-2000; 2000WO-US006792.
 XX
 PR 19-MAR-1999; 99US-0125362P.
 PR 10-DEC-1999; 99US-0169980P.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen GA, Ruben SM, Komatsoulis G;
 XX
 XX WPI; 2000-579483/54.
 DR N-PSDB; AAC74253.
 XX
 PT Isolated nucleic acid molecule encoding a human secreted protein is used
 PT in preventing, treating or ameliorating a medical condition.
 XX
 PS Disclosure; Page 52; 434pp; English.
 XX
 CC The polynucleotide sequences given in AAC74223-C74279 encode the human
 CC secreted proteins represented in AAB39179-B39226. Sequences AAB39227-
 CC B39308 are alternative proteins encoded by the genes, and also protein
 CC sequences with which they share homology. The proteins have activities
 CC based on the tissues and cells in which they are expressed. Examples of
 CC activities include: immunosuppressive; antiarthritic; antirheumatic;
 CC antiproliferative; cytostatic; cardiant; vasotropic; cerebroprotective;
 CC nootropic; neuroprotective; antibacterial; virucide; fungicide; and
 CC ophthalmological. The human secreted proteins, polynucleotides,
 CC antagonists and agonists of the invention may be useful in the treatment,
 CC prevention, and/or diagnosis of various disease, disorders and conditions
 CC such as autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative
 CC disorders e.g. neoplasms of the breast or liver, cardiovascular disorders
 CC e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia,
 CC angiogenesis, nervous system disorders e.g. Alzheimer's disease,
 CC infections caused by bacteria, viruses and fungi and ocular disorders
 CC e.g. corneal infection. The polypeptides can also be used to aid wound
 CC healing and epithelial cell proliferation, to regenerate tissues,
 CC maintain organs before transplantation, in chemotaxis and as a food
 CC additive or preservative e.g. to increase storage capabilities. Sequences
 CC AAC74214-C74222 and AAB39178 are used during the isolation and
 CC characterisation of the genes of the invention
 XX
 SQ Sequence 17 AA;
 Query Match 100.0%; Score 35; DB 3; Length 17;
 Best Local Similarity 100.0%; Pred. No. 71;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 2 PPPHG 6
 RESULT 4
 ABB70178
 ID ABB70178 standard; protein; 53 AA.
 XX
 AC ABB70178;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 37326.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US0009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR

PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL14281.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 37326; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 53 AA;
 Query Match 100.0%; Score 35; DB 4; Length 53;
 Best Local Similarity 100.0%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 2 PPPHG 6
 RESULT 5
 AAM15229
 ID AAM15229 standard; protein; 55 AA.
 XX
 AC AAM15229;
 XX
 DT 12-OCT-2001 (first entry)
 XX
 DE Peptide #1663 encoded by probe for measuring cervical gene expression.
 XX
 KW Probe; human; microarray; gene expression; cervical epithelial cell;
 KW cervical cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157278-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000670.
 XX
 XX 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 XX (MOLE-) MOLECULAR DYNAMICS INC.
 PA
 XX Penn SG, Hanzel DK, Chen W, Rank DR;
 PI
 XX WPI; 2001-488901/53.
 DR
 XX Human genome-derived single exon nucleic acid probes useful for analyzing

```

PT gene expression in human cervical epithelial cells.
PS Claim 27; SEQ ID NO 20055; 487pp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SENIP: see AA110068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENIPs are derived from human Hela cells. The SENIPs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical
CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer. Note: The sequence data for this patent did not form
CC part of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 55 AA;
SQ
    Query Match      100.0%; Score 35; DB 4; Length 55;
    Best Local Similarity 100.0%; Pred. No. 2.1e+02;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 PPPHG 5
    Db 1 PPPHG 5
    RESULT 6
    ID ABB34219 standard; peptide; 55 AA.
    XX
    AC ABB34219;
    DT 04-FEB-2002 (first entry)
    DE Peptide #1725 encoded by human foetal liver single exon probe.
    XX
    KW Human; foetal liver; gene expression; single exon nucleic acid probe.
    OS Homo sapiens.
    XX
    PN WO200157277-A2.
    PD 09-AUG-2001.
    PF 30-JAN-2001; 2001WO-US000669.
    PR 04-FEB-2000; 2000US-0180312P.
    PR 26-MAY-2000; 2000US-0207456P.
    PR 30-JUN-2000; 2000US-00608408.
    PR 03-AUG-2000; 2000US-00632366.
    PR 21-SEP-2000; 2000US-0234687P.
    PR 27-SEP-2000; 2000US-0236359P.
    PR 04-OCT-2000; 2000GB-00024263.
    XX
    PA (MOLE-) MOLECULAR DYNAMICS INC.
    XX
    PI Penn SG, Hanzel DK, Chen W, Rank DR;
    XX
    DR WPI; 2001-483447/52.
    XX
    PT Human genome-derived single exon nucleic acid probes useful for analyzing
    PT gene expression in human foetal liver.
    PS Claim 27; SEQ ID NO 26854; 639pp + Sequence Listing; English.
    XX
    CC The invention relates to a single exon nucleic acid probe for measuring
    CC human gene expression in a sample derived from human foetal liver. The
    CC single exon nucleic acid probes may be used for predicting, measuring and
    CC displaying gene expression in samples derived from human foetal liver. The
    CC present sequence is a peptide encoded by a single exon nucleic acid probe
    CC of the invention. Note: The sequence data for this patent did not form
    CC part of the printed specification, but was obtained in electronic format
    CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
    CC
    Query Match      100.0%; Score 35; DB 4; Length 55;
    Best Local Similarity 100.0%; Pred. No. 2.1e+02;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 PPPHG 5
    Db 1 PPPHG 5
    RESULT 7
    ID AAM27690 standard; protein; 55 AA.
    XX
    AC AAM27690;
    DT 17-OCT-2001 (first entry)
    DE Peptide #1727 encoded by probe for measuring placental gene expression.
    XX
    KW Probe; microarray; human; placenta; antenatal diagnosis;
    KW genetic disorder.
    XX
    OS Homo sapiens.
    XX
    PN WO200157272-A2.
    PD 09-AUG-2001.
    PF 30-JAN-2001; 2001WO-US000663.
    PR 04-FEB-2000; 2000US-0180312P.
    PR 26-MAY-2000; 2000US-0207456P.
    PR 30-JUN-2000; 2000US-00608408.
    PR 03-AUG-2000; 2000US-00632366.
    PR 21-SEP-2000; 2000US-0234687P.
    PR 27-SEP-2000; 2000US-0236359P.
    PR 04-OCT-2000; 2000GB-00024263.
    XX
    PA (MOLE-) MOLECULAR DYNAMICS INC.
    XX
    PI Penn SG, Hanzel DK, Chen W, Rank DR;
    XX
    DR WPI; 2001-488897/53.
    XX
    PT Human genome-derived single exon nucleic acid probes useful for analyzing
    PT gene expression in human placenta.
    PS Claim 27; SEQ ID NO 27959; 654pp; English.
    XX
    CC The present invention relates to single exon nucleic acid probes (SENIP:
    CC see AA13135-AA157546). The present sequence is a peptide encoded by one
    CC such probe. The probes are useful for producing a microarray for
    CC predicting, measuring and displaying gene expression in samples derived
    CC from human placenta. The probes are useful for antenatal diagnosis of
    CC human genetic disorders
    XX
    Query Match      100.0%; Score 35; DB 4; Length 55;
    Best Local Similarity 100.0%; Pred. No. 2.1e+02;
    Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
    QY 1 PPPHG 5
    Db 1 PPPHG 5
    RESULT 8
    ID ABB29058 standard; peptide; 55 AA.

```

```
XX ABB29058;
AC
XX
XX
XX 01-FEB-2002 (first entry)
XX
XX Peptide #1709 encoded by breast cell single exon nucleic acid probe.
DE
XX
XX Human; microarray; single exon probe; gene expression; breast; disease;
KW
KW cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200157271-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000662.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
PR
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-496933/54.
XX
XX New spatially-addressable set of single exon nucleic acid probes, useful
PT for measuring gene expression in sample derived from human breast,
PT comprises number of single exon nucleic acid probes.
XX
XX Claim 27; SEQ ID NO 12026; 327pp + Sequence Listing; English.
PS
XX
XX The invention relates to a spatially-addressable set of single exon
CC nucleic acid probes for measuring gene expression in a sample derived
CC from human breast and BT 474 cells. The method involves contacting the
CC probes with a collection of detectably labelled nucleic acids derived
CC from mRNA of human breast, and then measuring the label bound to each
CC probe of the microarray. The probes are useful for verifying the
CC expression of regions of genomic DNA predicted to encode proteins. They
CC are useful for gene discovery, and for determining predisposition and/or
CC prognosing breast disease. Gene expression analysis is useful for
CC assessing the toxicity of chemical agents on cells. The microarray of
CC this invention presents a far greater diversity of probes for measuring
CC gene expression, with far less bias than expressed sequence tag
CC microarrays. The method is suitable for rapid production of functional
CC information from genomic sequence. The present sequence is a peptide
CC encoded by a single exon nucleic acid probe of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 55 AA;
SQ
XX
XX Query Match 100.0%; Score 35; DB 4; Length 55;
XX Best Local Similarity 100.0%; Pred. No. 2.1e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PPPHG 5
XX |||||
XX Db 1 PPPHG 5
XX
XX RESULT 9
XX ABB19655
XX ID ABB19655 standard; protein; 55 AA.
XX
XX ABB19655;
AC
```

```
XX 23-JAN-2002 (first entry)
XX
XX Protein #1654 encoded by probe for measuring heart cell gene expression.
XX
XX Human; gene expression; heart; microarray; vascular system;
KW cardiovascular disease; hypertension; cardiac arrhythmia;
KW congenital heart disease.
XX
XX Homo sapiens.
OS
XX
XX WO200157274-A2.
PN
XX
XX 09-AUG-2001.
PD
XX
XX 30-JAN-2001; 2001WO-US000666.
PF
XX
XX 04-FEB-2000; 2000US-0180312P.
PR
XX 26-MAY-2000; 2000US-0207456P.
PR
XX 30-JUN-2000; 2000US-00608408.
PR
XX 03-AUG-2000; 2000US-00632366.
PR
XX 21-SEP-2000; 2000US-0234687P.
PR
XX 27-SEP-2000; 2000US-0236359P.
PR
XX 04-OCT-2000; 2000GB-00024263.
PR
XX (MOLE-) MOLECULAR DYNAMICS INC.
PA
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts.
PT
XX
XX Claim 15; SEQ ID NO 21425; 530pp; English.
PS
XX
XX The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21535-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 55 AA;
SQ
XX
XX Query Match 100.0%; Score 35; DB 4; Length 55;
XX Best Local Similarity 100.0%; Pred. No. 2.1e+02;
XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1 PPPHG 5
XX |||||
XX Db 1 PPPHG 5
XX
XX RESULT 10
XX AAM67398
XX ID AAM67398 standard; protein; 55 AA.
XX
XX AAM67398;
AC
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 27704.
DE
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KW microarray; cancer; leukaemia; lymphoma; myeloma.
XX
```

OS Homo sapiens.
 PN WO200157276-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000668.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488900/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.
 PT
 XX
 PS Example 4; SEQ ID NO 27704; 658pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC bone marrow. They can be used to measure gene expression in bone marrow
 CC samples, which may enable the improved diagnosis and treatment of cancers
 CC such as lymphoma, leukaemia and myeloma. The present sequence is a
 CC protein encoded by one of the probes of the invention
 CC
 XX Sequence 55 AA;
 SQ

Query Match 100.0%; Score 35; DB 4; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
 DB 1 PPHG 5

RESULT 11
 AAM55013
 ID AAM55013 standard; protein; 55 AA.
 XX
 AC AAM55013;
 XX
 DT 05-NOV-2001 (first entry)
 XX
 DE Human brain expressed single exon probe encoded protein SEQ ID NO: 27118.
 XX
 KW Human; brain expressed exon; gene expression analysis; probe; microarray;
 KW Alzheimer's disease; multiple sclerosis; schizophrenia; epilepsy; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200157275-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000667.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX

XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-483446/52.
 XX
 XX Single exon nucleic acid probes for analyzing gene expression in human
 PT brains.
 PT
 PS Example 4; SEQ ID NO 27118; 650pp + Sequence Listing; English.
 XX
 CC The present invention provides a number of single exon nucleic acid
 CC probes which are derived from genomic sequences expressed in the human
 CC brain. They can be used to measure gene expression in brain cell samples,
 CC which may enable the diagnosis and improved treatment of nervous system
 CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia, of
 CC epilepsy and cancers. The present sequence is a protein encoded by one of
 CC the probes of the invention
 XX
 SQ Sequence 55 AA;
 XX

Query Match 100.0%; Score 35; DB 4; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
 DB 1 PPHG 5

RESULT 12
 ABG49057
 ID ABG49057 standard; peptide; 55 AA.
 XX
 AC ABG49057;
 XX
 DT 25-FEB-2003 (first entry)
 XX
 DE Human liver peptide, SEQ ID NO 27705.
 XX
 KW Human; liver; cirrhosis; hyperlipoproteinaemia; hyperlipidaemia;
 KW hypercholesterolaemia; coronary heart disease.
 XX
 OS Homo sapiens.
 XX
 PN WO200157273-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 30-JAN-2001; 2001WO-US000664.
 XX
 PR 04-FEB-2000; 2000US-0180312P.
 PR 26-MAY-2000; 2000US-0207456P.
 PR 30-JUN-2000; 2000US-00608408.
 PR 03-AUG-2000; 2000US-00632366.
 PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488898/53.
 XX
 XX Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human adult liver.
 PT
 XX
 PS Claim 27; SEQ ID NO 27705; 658pp; English.
 XX
 XX The invention relates to a single exon nucleic acid probe (SEN) (1) for
 CC measuring human gene expression in a sample derived from human adult

CC liver, comprising one of 13109 defined nucleotide sequences given in the
CC specification (or complements/ fragments). The probe hybridises at high
CC stringency to a nucleic acid molecule expressed in the human adult liver.
CC (1) may be used for predicting, measuring and displaying gene expression
CC in samples derived from human adult liver. The genes identified may be
CC involved in genetic liver diseases such as cirrhosis,
CC hyperlipoproteinaemia, hyperlipidaemia and hypercholesterolaemia which is
CC associated with coronary heart disease. ABG47348-ABG59930 represent human
CC liver single exon encoded peptides of the invention. Note: The sequence
CC information for this patent does not appear in the printed specification
CC ftp.wipo.int/pub/published_pct_sequences
CC
XX
SQ Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPHG 5
Db 1 PPHG 5

RESULT 13

AA02972
ID AA02972 standard; protein; 55 AA.

XX
AC AA02972;
DT 09-OCT-2001 (first entry)

XX Peptide #1654 encoded by probe for measuring breast gene expression.
DE Probe; human; breast disease; breast cancer; development disorder;
XX inflammatory disease; proliferative breast disease; non-carcinoma tumour.
KW Homo sapiens.

OS
XX
XX WO200157270-A2.
FN 09-AUG-2001.

XX
XX
XX 29-JAN-2001; 2001WO-US000661.

XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-476286/51.

XX Novel single exon nucleic acid probe used to measuring gene expression in
XX a human breast.

XX Claim 27; SEQ ID NO 11712; 322pp; English.

XX The present invention relates to novel single exon nucleic acid probes
XX (see AA100010-AA110067). The present sequence is a peptide encoded by one
XX such probe. The probes are useful for measuring human gene expression in
XX a human breast sample, where the probe hybridises at high stringency to a
XX nucleic acid expressed in the human breast. The probes are useful for
XX predicting, diagnosing, grading, staging, monitoring and prognosing
XX diseases of the human breast, particularly those diseases with polygenic
XX aetiology. The diseases include: breast cancer, disorders of development,
XX inflammatory diseases of the breast, fibrocystic changes, proliferative

CC breast disease and non-carcinoma tumours. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 55 AA;

Query Match 100.0%; Score 35; DB 4; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPHG 5
Db 1 PPHG 5

RESULT 14

ABG37025
ID ABG37025 standard; peptide; 55 AA.

XX
AC ABG37025;

XX 19-AUG-2002 (first entry)

XX Human peptide encoded by genome-derived single exon probe SEQ ID 26690.

DE Human; single exon probe; asthma; lung cancer; COPD; ILD;
XX chronic obstructive pulmonary disease; interstitial lung disease;
KW familial idiopathic pulmonary fibrosis; neurofibromatosis;
KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;
KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemosiderosis;
KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;
KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;
KW primary ciliary dyskinesia; pulmonary hypertension;
KW hyaline membrane disease.

XX Homo sapiens.

OS
XX WO200186003-A2.

XX 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US000665.

XX 04-FEB-2000; 2000US-0180312P.
XX 26-MAY-2000; 2000US-0207456P.
XX 30-JUN-2000; 2000US-00608408.
XX 03-AUG-2000; 2000US-00632366.
XX 21-SEP-2000; 2000US-0234687P.
XX 27-SEP-2000; 2000US-0236359P.
XX 04-OCT-2000; 2000GB-00024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2002-114183/15.

XX Spatially-addressable set of single exon nucleic acid probes, used to
XX measure gene expression in human lung samples.

XX Claim 27; SEQ ID NO 26690; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon
XX nucleic acid probes for measuring gene expression in a sample derived
XX from human lung comprising single exon nucleic acid probes having one of
XX 12614 nucleic acid sequences mentioned in the specification, or their
XX complements or the 12387 open reading frames derived from the 12614
XX probes. Also included are a microarray comprising the novel set of probes
XX; the novel set of probes which hybridise at high stringency to a nucleic
XX acid expressed in the human lung; measuring gene expression in a sample
XX derived from human lung, comprising (a) contacting the array with a
XX collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of the
 CC array; identifying exons in a eukaryotic genome, comprising (a)
 CC algorithmically predicting at least one exon from genomic sequences of
 CC the eukaryote; and (b) detecting specific hybridisation of detectably
 CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,
 CC having a fragment identical to the predicted exon, the probe is included
 CC in the above mentioned microarray; assigning exons to a single gene,
 CC comprising (a) identifying exons from genomic sequence by the method
 CC above and (b) measuring the expression of each of the exons in several
 CC tissues and/or cell types using hybridisation to a single exon
 CC microarrays having a probe with the exon, where a common pattern of
 CC expression of the exons in the tissues and/or cell types indicates that
 CC the exons should be assigned to a single gene; a peptide comprising one
 CC of 12011 sequences, mentioned in the specification, or encoded by the
 CC probes/open reading frames (ORF). The probes are used for gene expression
 CC analysis, and for identifying exons in a gene, particularly using human
 CC lung derived mRNA and for the study of lung diseases such as asthma, lung
 CC cancer, chronic obstructive pulmonary disease (COPD), interstitial lung
 CC disease (ILD), familial idiopathic pulmonary fibrosis, neurofibromatosis,
 CC tuberous sclerosis, Gaucher's disease, Niemann-Pick disease, Hermansky-
 CC Pudlak syndrome, sarcoidosis, pulmonary haemosiderosis, pulmonary
 CC histiocytosis, lymphangioleiomyomatosis, pulmonary alveolar proteinosis,
 CC Karsagen syndrome, fibrocystic pulmonary dysplasia, primary ciliary
 CC dyskinesia, pulmonary hypertension and hyaline membrane disease. The
 CC present sequence is a peptide/protein encoded by a single exon probe of
 CC the invention. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 55 AA;

Query Match 100.0%; Score 35; DB 5; Length 55;
 Best Local Similarity 100.0%; Pred. No. 2.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
 DB 1 PPPHG 5

RESULT 15
 AAU54664
 ID AAU54664 standard; protein; 57 AA.

AC AAU54664;
 DT 27-FEB-2002 (first entry)
 XX
 DE Propionibacterium acnes immunogenic protein #15560.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

XX Propionibacterium acnes.

PN WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

PR 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;

PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59566.
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX Example 1; SEQ ID NO 15859; 1069pp; English.
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 XX Sequence 57 AA;

Query Match 100.0%; Score 35; DB 4; Length 57;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
 DB 47 PPPHG 51

RESULT 16
 ABM51183
 ID ABM51183 standard; protein; 57 AA.

AC ABM51183;

DT 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #15859.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX WPI; 2003-381789/36.

DR N-PSDB; ACF64495.

XX New Propionibacterium acnes polypeptides and polynucleotides encoding the

PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 15859; 1481pp; English.

XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM4536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridization. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 57 AA;

Query Match 100.0%; Score 35; DB 6; Length 57;
 Best Local Similarity 100.0%; Pred. No. 2.2e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;

QY 1 PPPHG 5
 |||||
 Db 47 PPPHG 51

RESULT 17
 AAU43160
 ID AAU43160 standard; protein; 60 AA.

AC AAU43160;

XX 27-FEB-2002 (first entry)

XX Propionibacterium acnes immunogenic protein #4056.

XX SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
 KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.

OS Propionibacterium acnes.

XX WO200181581-A2.

XX 01-NOV-2001.

XX 20-APR-2001; 2001WO-US012865.

XX 21-APR-2000; 2000US-0199047P.

PR 02-JUN-2000; 2000US-0208841P.

PR 07-JUL-2000; 2000US-0216747P.

XX (CORI-) CORIXA CORP.

XX

PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;

XX WPI; 2001-616774/71.

DR N-PSDB; AAS59519.

XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.

XX Example 1; SEQ ID NO 4355; 1069pp; English.

XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 60 AA;

Query Match 100.0%; Score 35; DB 4; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02; Mismatches 0; Indels 0; Gaps 0;
 Matches 5; Conservative 0;

QY 1 PPPHG 5
 |||||
 Db 46 PPPHG 50

RESULT 18
 ABM39679
 ID ABM39679 standard; protein; 60 AA.

AC ABM39679;

XX 20-OCT-2003 (first entry)

XX Propionibacterium acnes predicted ORF-encoded polypeptide #4355.

XX Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.

OS Propionibacterium acnes.

XX WO2003033515-A1.

XX 24-APR-2003.

XX 11-OCT-2002; 2002WO-US032727.

XX 15-OCT-2001; 2001US-00978825.

XX (CORI-) CORIXA CORP.

XX Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;

PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;

PI Barth B, Vallieve-Douglass J;

XX

DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64448.
 XX
 PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 4355; 1481pp; English.
 XX
 CC The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC polynucleotide of the invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 60 AA;

Query Match 100.0%; Score 35; DB 6; Length 60;
 Best Local Similarity 100.0%; Pred. No. 2.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
 Db 46 PPPHG 50
 |||||

RESULT 19
 AAM37888
 ID AAM37888 standard; protein; 61 AA.

XX AC AAM37888;

XX DT 17-OCT-2001 (first entry)

XX DE Peptide #11925 encoded by probe for measuring placental gene expression.

XX KW Probe; microarray; human; placenta; antenatal diagnosis;
 XX genetic disorder.

XX OS Homo sapiens.

XX PN WO200157272-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000663.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

PR 21-SEP-2000; 2000US-0234687P.
 PR 27-SEP-2000; 2000US-0236359P.
 PR 04-OCT-2000; 2000GB-00024263.
 XX
 PA (MOLE-) MOLECULAR DYNAMICS INC.
 XX
 PI Penn SG, Hanzel DK, Chen W, Rank DR;
 XX
 DR WPI; 2001-488897/53.
 XX
 PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human placenta.
 XX
 PS Claim 27; SEQ ID NO 38157; 654pp; English.
 XX
 CC The present invention relates to single exon nucleic acid probes (SENP:
 CC see AA13115-AA157546). The present sequence is a peptide encoded by one
 CC such probe. The probes are useful for producing a microarray for
 CC predicting, measuring and displaying gene expression in samples derived
 CC from human placenta. The probes are useful for antenatal diagnosis of
 CC human genetic disorders
 XX
 SQ Sequence 61 AA;

Query Match 100.0%; Score 35; DB 4; Length 61;

Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 18 PPPHG 22
 |||||

RESULT 20

AAM77682
 ID AAM77682 standard; protein; 61 AA.

XX AC AAM77682;

XX DT 06-NOV-2001 (first entry)

XX DE Human bone marrow expressed probe encoded protein SEQ ID NO: 37988.

XX KW Human; bone marrow expressed exon; gene expression analysis; probe;
 XX microarray; cancer; leukaemia; lymphoma; myeloma.

XX OS Homo sapiens.

XX PN WO200157276-A2.

XX PD 09-AUG-2001.

XX PF 30-JAN-2001; 2001WO-US000668.

XX PR 04-FEB-2000; 2000US-0180312P.

XX PR 26-MAY-2000; 2000US-0207456P.

XX PR 30-JUN-2000; 2000US-00608408.

XX PR 03-AUG-2000; 2000US-00632366.

XX PR 21-SEP-2000; 2000US-0234687P.

XX PR 27-SEP-2000; 2000US-0236359P.

XX PR 04-OCT-2000; 2000GB-00024263.

XX PA (MOLE-) MOLECULAR DYNAMICS INC.

XX PI Penn SG, Hanzel DK, Chen W, Rank DR;

XX DR WPI; 2001-488900/53.

XX PT Human genome-derived single exon nucleic acid probes useful for analyzing
 PT gene expression in human bone marrow.

XX PS Example 4; SEQ ID NO 37988; 658pp + Sequence Listing; English.

XX

XX AC AAU32380;
 XX DT 18-DEC-2001 (first entry)
 XX DE Novel human secreted protein #2871.
 XX XX Human; vaccination; gene therapy; nutritional supplement;
 KW stem cell proliferation; haematopoiesis; nerve tissue regeneration;
 KW immune suppression; immune stimulation; anti-inflammatory; leukaemia.
 XX OS Homo sapiens.
 XX PN WO200179449-A2.
 XX PD 25-OCT-2001.
 XX PF 16-APR-2001; 2001WO-US008656.
 XX PR 18-APR-2000; 2000US-00552929.
 XX PR 26-JAN-2001; 2001US-00770160.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang YT, Liu C, Drmanac RT;
 XX DR WPI; 2001-611725/70.
 XX PT Nucleic acids encoding a range of human polypeptides, useful in genetic
 PT vaccination, testing and therapy.
 XX PS Claim 20; Page 597-598; 765pp; English.
 CC The invention relates to novel human secreted polypeptides. The
 CC polypeptides and antibodies to the polypeptides are useful for
 CC determining the presence of or predisposition to a disease associated
 CC with altered levels of polypeptide. The polypeptides are also useful for
 CC identifying agents (agonists and antagonists) that bind to them. Cells
 CC expressing the proteins are useful for identifying a therapeutic agent
 CC for use in treatment of a pathology related to aberrant expression or
 CC physiological interactions of the polypeptide. Vectors comprising the
 CC nucleic acids encoding the polypeptides and cells genetically engineered
 CC to express them are also useful for producing the proteins. The proteins
 CC are useful in genetic vaccination, testing and therapy, and can be used
 CC as nutritional supplements. They may be used to increase stem cell
 CC proliferation; to regulate haematopoiesis; and in bone, cartilage, tendon
 CC and/or nerve tissue growth or regeneration; immune suppression and/or
 CC stimulation; as anti-inflammatory agents; and in treatment of leukaemias.
 CC AAU29510-AAU3304 represent the amino acid sequences of novel human
 CC secreted proteins of the invention
 XX SQ Sequence 75 AA;
 Query Match 100.0%; Score 35; DB 4; Length 75;
 Best Local Similarity 100.0%; Pred. NO. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPHG 5
 Db 3 PPHG 7
 RESULT 24
 AAU41116
 ID AAU41116 standard; protein; 77 AA.
 XX AC AAU41116;
 XX DT 13-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #2012.
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 XX dermatological; osteopathic; neuroprotectant.
 OS Propionibacterium acnes.
 XX PN WO200181581-A2.
 XX PD 01-NOV-2001.
 XX PF 20-APR-2001; 2001WO-US012865.
 XX PR 21-APR-2000; 2000US-0199047P.
 XX PR 02-JUN-2000; 2000US-0208841P.
 XX PR 07-JUL-2000; 2000US-0216747P.
 XX PA (CORI-) CORIXA CORP.
 XX PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX DR WPI; 2001-616774/71.
 XX DR N-FSDB; AAS59514.
 XX PT Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 XX PS Example 1; SEQ ID NO 2311; 1069pp; English.
 XX CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hypertosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 77 AA;
 Query Match 100.0%; Score 35; DB 4; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPHG 5
 Db 27 PPHG 31
 RESULT 25
 AAU43934
 ID AAU43934 standard; protein; 77 AA.
 XX AC AAU43934;
 XX DT 27-FEB-2002 (first entry)
 XX DE Propionibacterium acnes immunogenic protein #4830.
 KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;

KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
 KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
 KW dermatological; osteopathic; neuroprotectant.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO200181581-A2.
 XX
 XX 01-NOV-2001.
 PD
 XX
 PF 20-APR-2001; 2001WO-US012865.
 XX
 XX 21-APR-2000; 2000US-0199047P.
 PR 02-JUN-2000; 2000US-0208841P.
 PR 07-JUL-2000; 2000US-0216747P.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 XX Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
 PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
 XX
 DR WPI; 2001-616774/71.
 DR N-PSDB; AAS59521.
 XX
 XX Propionibacterium acnes polypeptides and nucleic acids useful for
 PT vaccinating against and diagnosing infections, especially useful for
 PT treating acne vulgaris.
 PT
 XX
 PS Example 1; SEQ ID NO 5129; 1069pp; English.
 XX
 XX Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
 CC polypeptides. The proteins and their associated DNA sequences are used in
 CC the treatment, prevention and diagnosis of medical conditions caused by
 CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
 CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
 CC P. acnes is also involved in infections of bone, joints and the central
 CC nervous system, however it is particularly involved in the inflammatory
 CC lesions associated with acne vulgaris. A method for detecting the
 CC presence or absence of P. acnes in a patient comprises contacting a
 CC sample with a binding agent that binds to the proteins of the invention
 CC and determining the amount of bound protein in the sample. The
 CC polypeptides may be used as antigens in the production of antibodies
 CC specific for P. acnes proteins. These antibodies can be used to
 CC downregulate expression and activity of P. acnes polypeptides and
 CC therefore treat P. acnes infections. The antibodies may also be used as
 CC diagnostic agents for determining P. acnes presence, for example, by
 CC enzyme linked immunosorbent assay (ELISA). Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 77 AA;
 Query Match 100.0%; Score 35; DB 4; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 13 PPPHG 17
 RESULT 26
 ABM40453
 ID ABM40453 standard; protein; 77 AA.
 XX
 AC ABM40453;
 XX
 XX 20-OCT-2003 (first entry)
 DT
 DE Propionibacterium acnes predicted ORF-encoded polypeptide #5129.
 XX
 XX Acne vulgaris; antisecborrhoeic; dermatological; antibacterial;
 KW

KW immunostimulant; immune response; vaccine.
 XX
 OS Propionibacterium acnes.
 XX
 PN WO2003033515-A1.
 XX
 PD 24-APR-2003.
 XX
 PF 11-OCT-2002; 2002WO-US032727.
 XX
 PR 15-OCT-2001; 2001US-00978825.
 XX
 XX (CORI-) CORIXA CORP.
 PA
 PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglas J;
 XX
 DR WPI; 2003-381789/36.
 DR N-PSDB; ACF64450.
 XX
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX
 PS Example 1; SEQ ID NO 5129; 1481pp; English.
 XX
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising P. acnes prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 77 AA;
 Query Match 100.0%; Score 35; DB 6; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 13 PPPHG 17
 RESULT 27
 ABM37635
 ID ABM37635 standard; protein; 77 AA.
 XX
 AC ABM37635;
 XX
 XX 20-OCT-2003 (first entry)
 DT

XX DE Propionibacterium acnes predicted ORF-encoded polypeptide #2311.
 XX KW Acne vulgaris; antiseborrheic; dermatological; antibacterial;
 KW immunostimulant; immune response; vaccine.
 XX OS Propionibacterium acnes.
 XX PN WO2003033515-A1.
 XX PD 24-APR-2003.
 XX PF 11-OCT-2002; 2002WO-US032727.
 XX PR 15-OCT-2001; 2001US-00978825.
 XX PA (CORI-) CORIXA CORP.
 XX PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
 PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
 PI Barth B, Vallieue-Douglass J;
 XX WPI: 2003-381789/36.
 DR N-PSDB; ACF64443.
 XX New Propionibacterium acnes polypeptides and polynucleotides encoding the
 PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
 PT or for stimulating an immune response specific for a P. acnes protein.
 XX Example 1; SEQ ID NO 2311; 1481pp; English.
 XX The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
 CC encoding a Propionibacterium acnes protein. The invention also relates to
 CC polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
 CC immunogenic fragments of P. acnes polypeptides. The invention
 CC additionally encompasses expression vectors and host cells comprising a
 CC polynucleotide of the invention; antibodies against polypeptides of the
 CC invention; fusion proteins comprising a polypeptide of the invention; a
 CC method for stimulating an immune response specific for a P. acnes
 CC polypeptide and an isolated T cell population comprising T cells prepared
 CC via this method; a vaccine composition (comprising P. acnes polypeptides,
 CC polynucleotides, antibodies, fusion proteins, T cell populations, or
 CC antigen-presenting cells that express the polypeptide); a method and kit
 CC for detecting or determining the presence or absence of P. acnes in a
 CC patient; and a method for inhibiting the development of P. acnes in a
 CC patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
 CC proteins, T cell populations or antigen-presenting cells that express the
 CC polypeptides are useful for diagnosing, preventing or treating acne
 CC vulgaris, or for stimulating an immune response specific for a P. acnes
 CC protein. The polynucleotides can also be used as probes or primers for
 CC nucleic acid hybridisation. The vaccine composition is useful for the
 CC stimulation of an immune response against P. acnes, or for treating acne,
 CC and the kit is useful for performing a diagnostic assay. The present
 CC sequence represents a polypeptide predicted to be encoded by an ORF (open
 CC reading frame) contained within the P. acnes polynucleotides of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX SQ Sequence 77 AA;
 Query Match 100.0%; Score 35; DB 6; Length 77;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 27 PPPHG 31
 RESULT 28
 AAY12827
 ID AAY12827 standard; protein; 79 AA.

XX AAY12827;
 XX 21-JUN-1999 (first entry)
 XX Human 5' EST secreted protein SEQ ID NO:417.
 XX Human; secreted protein; EST; expressed sequence tag; diagnosis;
 KW forensic; gene therapy; chromosome mapping; signal peptide;
 KW upstream regulatory sequence; cytokine activity; cell proliferation;
 KW differentiation; haematopoiesis regulation; tissue growth regulation;
 KW reproductiv hormone regulation; chemotactic; chemokinetic; haemostatic;
 KW thrombolytic; anti-inflammatory; tumour inhibition.
 XX Homo sapiens.
 XX WO9906549-A2.
 XX 11-FEB-1999.
 XX 31-JUL-1998; 98WO-IB001231.
 XX 01-AUG-1997; 97US-00905279.
 XX (GEST) GENSET.
 XX Dumas Milne Edwards J, Duclert A, Lacroix B;
 XX WPI; 1999-153779/13.
 DR N-PSDB; AAX51605.
 XX New nucleic acids encoding human secreted proteins - obtained from cDNA
 PT libraries derived from testis, ovary, uterus and spleen tissue.
 XX Claim 34; Page 469; 522pp; English.
 XX AAX51459 to AAX51691 represent 5' expressed sequence tags (ESTs) for
 CC human secreted proteins, and encode the proteins given in AAY12681 to
 CC AAY12913, respectively. The proteins given represent the signal peptide
 CC and an N-terminal fragment of a secreted protein. The nucleic acid
 CC sequences can be used for producing secreted human gene products. They
 CC can also be used to develop products for diagnosis and therapy. The
 CC proteins obtained may have cytokine activity, cell
 CC proliferation/differentiation activity, haematopoiesis regulating
 CC activity, tissue growth regulating activity, chemokinetic activity, haemostatic and
 CC thrombolytic activity, chemotactic/ chemokinetic activity, anti-inflammatory
 CC activity, tumour inhibition activity or other activities. The products
 CC can be used in forensic, gene therapy and chromosome mapping procedures.
 CC The sequences can also be used for obtaining corresponding promoter
 CC sequences. The nucleic acids encoding the signal peptide can be used for
 CC directing extracellular secretion of a polypeptide or the insertion of a
 CC polypeptide into a membrane, or importing a polypeptide into a cell
 XX SQ Sequence 79 AA;
 Query Match 100.0%; Score 35; DB 2; Length 79;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 59 PPPHG 63
 RESULT 29
 ABB61183
 ID ABB61183 standard; protein; 80 AA.
 XX ABB61183;
 XX 26-MAR-2002 (first entry)
 XX

DE Drosophila melanogaster polypeptide SEQ ID NO 10341.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN W0200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 XX 23-MAR-2001; 2001WO-US009231.
 PF
 XX 23-MAR-2000; 2000US-0191637P.
 PR
 PR 11-JUL-2000; 2000US-00614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI; 2001-656860/75.
 DR N-PSDB; ABL05286.
 XX
 XX New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signaling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 10341; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 80 AA;
 Query Match 100.0%; Score 35; DB 4; Length 80;
 Best Local Similarity 100.0%; Pred. No. 3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 12 PPPHG 16
 |||||
 RESULT 30
 ABP11024
 ID ABP11024 standard; protein; 89 AA.
 XX
 AC ABP11024;
 XX
 DT 25-JUN-2002 (first entry)
 XX
 DE Human ORFX protein sequence SEQ ID NO:22030.
 XX
 KW Human; open reading frame; ORFX; gene therapy; cancer; cirrhosis;
 KW hyperproliferative disorder; psoriasis; benign tumour; haemorrhage;
 KW degenerative disorder; osteoarthritis; neurodegenerative disorder;
 KW cardiovascular disease; diabetes mellitus; systemic lupus erythematosus;
 KW hypertension; hypothyroidism; cholesterol ester storage disease;
 KW immune deficiency; immune disorder; infectious disease;
 KW autoimmune disorder; rheumatoid arthritis; autoimmune thyroiditis;
 KW myasthenia gravis.
 XX
 OS Homo sapiens.
 XX
 PN W0200192523-A2.

XX 06-DEC-2001.
 PD
 XX 29-MAY-2001; 2001WO-US010836.
 PF
 XX 30-MAY-2000; 2000US-0206132P.
 PR
 PR 29-AUG-2000; 2000US-0228716P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX
 XX Shimkets RA, Leach MD;
 PI
 XX WPI; 2002-106308/14.
 DR N-PSDB; ABL26776.
 XX
 XX Novel human polypeptides and polynucleotides useful for diagnosing,
 PT preventing and treating cardiovascular disease, neurodegenerative,
 PT hyperproliferative disorders and autoimmune disorders.
 XX
 PS Disclosure; SEQ ID NO 22030; 1037pp; English.
 XX
 CC The present invention describes substantially purified human proteins
 CC (referred to as open reading frame, ORFX, where X is 1-11491 (see Table 1
 CC in the specification). ABL15762 to ABL27252 encode the human ORFX
 CC proteins given in ABL00010 to ABL1500. ORFX proteins are useful for
 CC treating or preventing a pathology associated with an ORFX-associated
 CC disorder in humans, and in the manufacture of a medicament for treating a
 CC syndrome associated with ORFX-associated disorder. ORFX polynucleotide
 CC sequences can be used in gene therapy. ORFX sequences can be used in the
 CC treatment of cancer, hyperproliferative disorders, cirrhosis of liver,
 CC psoriasis, benign tumours, keloid, degenerative disorders, haemorrhage,
 CC osteoarthritis, neurodegenerative disorders, disorders related to organ
 CC transplantation, cardiovascular diseases, diabetes mellitus, systemic
 CC lupus erythematosus, hypertension, hypothyroidism, cholesterol ester
 CC storage disease, various immune deficiencies and disorders, infectious
 CC disease, autoimmune disorders such as multiple sclerosis, rheumatoid
 CC arthritis, autoimmune thyroiditis, myasthenia gravis, graft-versus-host
 CC disease and autoimmune inflammatory eye disease. ORFX proteins are also
 CC useful for treating burns, incisions, ulcers, for treating osteoporosis,
 CC bone degenerative disorders, or periodontal disease, and for gut
 CC protection or regeneration and treatment of lung or liver fibrosis,
 CC reperfusion injury in various tissues and conditions resulting from
 CC systemic cytokine damage. N.B. The sequence data for this patent did not
 CC form part of the printed specification, but was obtained in electronic
 CC format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 89 AA;
 Query Match 100.0%; Score 35; DB 5; Length 89;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 68 PPPHG 72
 |||||
 RESULT 31
 ABU34317
 ID ABU34317 standard; protein; 89 AA.
 XX
 AC ABU34317;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #19844.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 KW Mycobacterium avium.
 XX
 OS
 XX
 PN W0200277183-A2.
 XX

PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 25-OCT-2001; 2001US-0342923F.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA38187.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 62241; 1766pp; English.
 XX
 CC The invention relates to an isolated nucleic acid comprising any one of
 CC the 6213 antisense sequences given in the specification where expression
 CC of the nucleic acid inhibits proliferation of a cell. Also included are:
 CC (1) a vector comprising a promoter operably linked to the nucleic acid
 CC encoding a polypeptide whose expression is inhibited by the antisense
 CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
 CC polypeptide or its fragment whose expression is inhibited by the
 CC antisense nucleic acid; (4) an antibody capable of specifically binding
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 CC proliferation or the activity of a gene in an operon required for
 CC proliferation; (7) identifying a compound that influences the activity of
 CC the gene product or that has an activity against a biological pathway
 CC required for proliferation, or that inhibits cellular proliferation; (8)
 CC identifying a gene required for cellular proliferation or the biological
 CC pathway in which a proliferation-required gene or its gene product lies
 CC or a gene on which the test compound that inhibits proliferation of an
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
 CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 89 AA;
 Query Match 100.0%; Score 35; DB 6; Length 89;
 Best Local Similarity 100.0%; Pred. No. 3.3e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 82 PPPHG 86
 RESULT 32
 ABG24494
 ID ABG24494 standard; protein; 97 AA.
 XX
 AC ABG24494;
 XX

DT 18-FEB-2002 (first entry)
 XX
 DE Novel human diagnostic protein #24485.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 KW food supplement; medical imaging; diagnostic; genetic disorder.
 OS
 OS Homo sapiens.
 XX
 FN WO200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 XX (HYSE-) HYSEQ INC.
 PA
 PA Drmanac RT, Liu C, Tang YT;
 PI
 XX
 DR WPI; 2001-639362/73.
 DR N-PSDB; AAS88681.
 XX
 CC New isolated polynucleotide and encoded polypeptides, useful in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits and to assess
 CC biodiversity.
 XX
 PS Claim 20; SEQ ID NO 54853; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. ABG00010-ABG30377 represent novel human diagnostic
 CC amino acid sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 97 AA;
 Query Match 100.0%; Score 35; DB 4; Length 97;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db |||||
 51 PPPHG 55
 RESULT 33
 ABO63515
 ID ABO63515 standard; protein; 98 AA.
 XX
 AC ABO63515;
 XX
 DT 29-JUL-2004 (first entry)
 XX

DE Klebsiella pneumoniae polypeptide seqid 10032.
 XX Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 XX Klebsiella pneumoniae.
 OS
 PN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX
 PA (GENO-) GENOME THERAPEUTICS CORP.
 XX
 PI Breton GL, Osborne M;
 DR WPI: 2003-895346/82.
 DR N-PSDB; ACH97066.
 XX
 XX New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for
 PT preparing a vaccine composition against Klebsiella pneumoniae.
 PT
 XX Disclosure; SEQ ID NO 10032; 932pp; English.
 PS
 CC The invention describes a new isolated nucleic acid encoding a Klebsiella
 CC pneumoniae polypeptide. Also described are: a recombinant expression
 CC vector comprising the nucleic acid, operably linked to a transcription
 CC regulatory element; and a cell comprising the recombinant expression
 CC vector. The nucleic acid is useful for preparing a vaccine composition
 CC against Klebsiella pneumoniae. This is the amino acid sequence of a
 CC Klebsiella pneumoniae polypeptide of the invention
 XX
 SQ Sequence 98 AA;
 Query Match 100.0%; Score 35; DB 7; Length 98;
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PPPHG 5
 Db 28 PPPHG 32
 RESULT 34
 ADH10412
 ID ADH10412 standard; protein; 101 AA.
 XX
 AC ADH10412;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Rabbit HPRG protein H/P rich domain.
 XX
 KW Tpm; tropomyosin; antiangiogenic receptor;
 KW histidine-proline rich glycoprotein; HPRG; cytostatic; antidiabetic;
 KW ophthalmological; antiinflammatory; gynaecological; antiarthritic;
 KW antipsoriatic; dermatological; cardiac; vasotropic; vulnery;
 KW angiogenesis; gene therapy; rabbit.
 XX
 OS Oryctolagus cuniculus.
 XX
 PN WO2003077872-A2.
 XX
 PD 25-SEP-2003.
 XX
 PF 17-MAR-2003; 2003WO-US008060.
 XX
 PR 15-MAR-2002; 2002US-0364047P.
 XX
 PA (ATTE-) ATTENUON LLC.

XX
 PI Mcrae K, Donate F, Juarez J, Mazar AP;
 XX WPI: 2004-090604/09.
 DR
 XX
 PT New tropomyosin-related antiangiogenic receptor polypeptide, useful for
 PT inhibiting endothelial cell migration, invasion, proliferation or
 PT angiogenesis, inducing endothelial cell apoptosis, or treating tumors or
 PT cancer.
 XX
 PS Claim 6; SEQ ID NO 24; 117pp; English.
 XX
 CC The invention relates to an isolated tropomyosin (Tpm)-related
 CC antiangiogenic receptor polypeptide or peptide, which is a fragment
 CC of a full-length native Tpm protein expressed on the surface of
 CC endothelial cells, or a variant of the fragment. It has a molecular mass
 CC of about 17 kDa and corresponds in its sequence to, or is a variant of,
 CC an internal fragment of a native Tpm isoform which is a binding site for
 CC antiangiogenic polypeptide agents; The isolated antiangiogenic receptor
 CC polypeptide, peptide or variant has substantially the same biochemical
 CC activity of binding to the antiangiogenic polypeptide agents, as does the
 CC native Tpm internal fragment. The antiangiogenic polypeptide agent that
 CC binds to the isolated polypeptide or peptide is human histidine-proline
 CC rich glycoprotein (HPRG), rabbit HPRG, a Tpm-binding, antiangiogenic
 CC homologue, variant, domain or fragment of human or rabbit HPRG, two chain
 CC human kininogen (HK), the D5 domain of HK, or a Tpm-binding,
 CC antiangiogenic homologue, variant, domain or fragment of the HK or its D5
 CC domain. The Tpm-related antiangiogenic receptor polypeptide or peptide,
 CC antibodies and compositions are useful for inhibiting endothelial cell
 CC migration, invasion, proliferation or angiogenesis, for inducing
 CC endothelial cell apoptosis, or for treating tumors or cancer, diabetic
 CC retinopathy, neovascular glaucoma, uveitis, endometriosis, arthritis,
 CC psoriasis, or scleroderma. The antibody may be also used for detecting
 CC the presence of a Tpm polypeptide or peptide in a biological sample, for
 CC promoting wound healing, or for treating diseases or conditions in which
 CC increased angiogenesis is desired, e.g. coronary artery disease or
 CC peripheral artery disease. The present sequence represents a rabbit HPRG
 CC protein His-Pro (H/P) rich domain.
 XX
 SQ Sequence 101 AA;
 Query Match 100.0%; Score 35; DB 8; Length 101;
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1 PPPHG 5
 Db 8 PPPHG 12
 RESULT 35
 ABO66179
 ID ABO66179 standard; protein; 103 AA.
 XX
 AC ABO66179;
 XX
 DT 29-JUL-2004 (first entry)
 XX
 DE Klebsiella pneumoniae polypeptide seqid 12696.
 XX
 KW Recombinant expression vector; transcription regulatory element;
 KW Klebsiella pneumoniae protein; antibacterial; Vaccine.
 XX
 OS Klebsiella pneumoniae.
 XX
 PN US6610836-B1.
 XX
 PD 26-AUG-2003.
 XX
 PF 27-JAN-2000; 2000US-00489039.
 XX
 PR 29-JAN-1999; 99US-0117747P.
 XX

PA	(GENO-) GENOME THERAPEUTICS CORP.	CC	treatment effect. Glycosylhydrolase 12 has cytostatic, antiviral, anti-
XX	Breton GL, Osborne M;	CC	HIV, immunomodulatory and antiinflammatory activities. The
XX	DR WPI; 2003-895346/82.	CC	glycosylhydrolase protein and polynucleotide sequence can be used for
XX	DR N-PSDB; ACH99730.	CC	treating various diseases, such as malignant tumour, haemopathy, HIV
XX		XX	Infection, immunological diseases and various inflammations
PT	New nucleic acid encoding a Klebsiella pneumoniae polypeptide, useful for	SQ	Sequence 111 AA;
PT	preparing a vaccine composition against Klebsiella pneumoniae.		
PS	Disclosure; SEQ ID NO 12696; 932pp; English.		
XX			
CC	The invention describes a new isolated nucleic acid encoding a Klebsiella		
CC	pneumoniae polypeptide. Also described are: a recombinant expression		
CC	vector comprising the nucleic acid, operably linked to a transcription		
CC	regulatory element; and a cell comprising the recombinant expression		
CC	vector. The nucleic acid is useful for preparing a vaccine composition		
CC	against Klebsiella pneumoniae. This is the amino acid sequence of a		
CC	Klebsiella pneumoniae polypeptide of the invention		
XX			
SQ	Sequence 103 AA;		
	Query Match 100.0%; Score 35; DB 7; Length 103;		Query Match 100.0%; Score 35; DB 4; Length 111;
	Best Local Similarity 100.0%; Pred. No. 3.8e+02;		Best Local Similarity 100.0%; Pred. No. 4e+02;
	Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 PPPHG 5	QY	1 PPPHG 5
DB		DB	
	10 PPPHG 14		53 PPPHG 57
RESULT 36		RESULT 37	
AAG67890		AAG64349	
ID	AAG67890 standard; protein; 111 AA.	ID	AAG64349 standard; protein; 111 AA.
XX		XX	
AC	AAG67890;	AC	AAG64349;
XX		XX	
DT	07-DEC-2001 (first entry)	DT	25-SEP-2001 (first entry)
XX		XX	
DE	Human glycosylhydrolase 12 protein SEQ ID NO:2.	DE	Glycosyl hydrolase.
XX		XX	
KW	Human; glycosylhydrolase 12; cytostatic; antiviral; anti-HIV;	KW	Glycosyl hydrolase; cytostatic; haemostatic; virucide; immunomodulatory;
KW	immunomodulatory; antiinflammatory; malignant tumour; haemopathy;	KW	antiinflammatory; gene therapy; malignant tumour; haemopathy;
KW	HIV Infection; immunological disease; inflammation.	KW	HIV infection; immunological disease; inflammation.
XX		XX	
OS	Homo sapiens.	OS	Unidentified.
XX		XX	
PN	CN1301842-A.	PN	WO200147991-A1.
XX		XX	
PD	04-JUL-2001.	PD	05-JUL-2001.
XX		XX	
PF	27-DEC-1999; 99CN-00125788.	PF	18-DEC-2000; 2000WO-CN000635.
XX		XX	
PR	27-DEC-1999; 99CN-00125788.	PR	24-DEC-1999; 99CN-00125774.
XX		XX	
PA	(BODE-) BODE GENE DEV CO LTD SHANGHAI.	PA	(BIOW-) BIOWINDOW GENE DEV INC SHANGHAI.
XX		XX	
PI	Mao Y, Xie Y;	PI	Mao Y, Xie Y;
XX		XX	
WPI	2001-550479/62.	WPI	2001-425642/45.
DR	N-PSDB; AAH44692.	DR	N-PSDB; AAH49695.
XX		XX	
XX	New glycosylhydrolase 12 polypeptide and the polynucleotide encoding it,	XX	Glycosyl hydrolase and encoded polynucleotide, used in diagnosis and
PT	useful in the treatment of various diseases, such as malignant tumor,	PT	treatment of malignant tumors, hemopathy, human immunodeficiency virus
PT	hemopathy, HIV Infection, immunological diseases and various	PT	infection, immunological diseases and inflammation.
PT	inflammations.	XX	
XX		PS	Claim 1; Page 28; 34pp; Chinese.
PS	Claim 1; Page 24 (Disclosure); 30pp; Chinese.	XX	
XX		XX	
CC	The present sequence represents the human glycosylhydrolase 12 protein.	XX	The present sequence is the protein sequence for glycosyl hydrolase. The
CC	Also described are: (1) polynucleotides encoding for the	CC	hydrolase and its coding sequence are useful in the diagnosis and
CC	glycosylhydrolase 12 protein; (2) production of the protein using a DNA	CC	treatment of malignant tumour, haemopathy, HIV infection, immunological
CC	recombination process; and (3) an agonist resisting the protein and its	CC	diseases and various inflammations

XX AAG16968; 99US-0139462P.
AC 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
DT 17-OCT-2000 (first entry) 99US-0139763P.
XX 18-JUN-1999; 99US-0139817P.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 17811. 99US-0139899P.
XX 22-JUN-1999; 99US-0140353P.
DE Protein identification; signal transduction pathway; metabolic pathway; 99US-0140354P.
KW hybridisation assay; genetic mapping; gene expression control; promoter; 99US-0140695P.
KW termination sequence. 99US-0140823P.
XX 28-JUN-1999; 99US-0140991P.
XX Arabidopsis thaliana. 99US-0141287P.
XX 30-JUN-1999; 99US-0141842P.
XX EP1033405-A2. 99US-0142154P.
XX 06-SEP-2000. 99US-0142055P.
XX 02-JUL-1999; 99US-0142390P.
XX 06-JUL-1999; 99US-0142803P.
XX 08-JUL-1999; 99US-0142920P.
XX 09-JUL-1999; 99US-0142977P.
XX 12-JUL-1999; 99US-0143542P.
XX 13-JUL-1999; 99US-0143624P.
XX 14-JUL-1999; 99US-0144005P.
XX 15-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 20-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145086P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
XX 22-JUL-1999; 99US-0145089P.
XX 22-JUL-1999; 99US-0145192P.
XX 23-JUL-1999; 99US-0145145P.
XX 23-JUL-1999; 99US-0145218P.
XX 23-JUL-1999; 99US-0145224P.
XX 26-JUL-1999; 99US-0145276P.
XX 27-JUL-1999; 99US-0145913P.
XX 27-JUL-1999; 99US-0145918P.
XX 27-JUL-1999; 99US-0145918P.
XX 28-JUL-1999; 99US-0145919P.
XX 28-JUL-1999; 99US-0145951P.
XX 02-AUG-1999; 99US-0146386P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146388P.
XX 02-AUG-1999; 99US-0146389P.
XX 03-AUG-1999; 99US-0147038P.
XX 04-AUG-1999; 99US-0147204P.
XX 04-AUG-1999; 99US-0147302P.
XX 05-AUG-1999; 99US-0147192P.
XX 05-AUG-1999; 99US-0147260P.
XX 06-AUG-1999; 99US-0147303P.
XX 06-AUG-1999; 99US-0147416P.
XX 09-AUG-1999; 99US-0147493P.
XX 09-AUG-1999; 99US-0147935P.
XX 10-AUG-1999; 99US-0148171P.
XX 11-AUG-1999; 99US-0148319P.
XX 12-AUG-1999; 99US-0148341P.
XX 13-AUG-1999; 99US-0148565P.
XX 13-AUG-1999; 99US-0148684P.
XX 16-AUG-1999; 99US-0149368P.
XX 17-AUG-1999; 99US-0149175P.
XX 18-AUG-1999; 99US-0149426P.
XX 20-AUG-1999; 99US-0149722P.
XX 20-AUG-1999; 99US-0149723P.
XX 20-AUG-1999; 99US-0149929P.
XX 23-AUG-1999; 99US-0149902P.
XX 23-AUG-1999; 99US-0149930P.
XX 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 23-APR-1999; 99US-0130891P.
XX 28-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 30-APR-1999; 99US-0132407P.
XX 04-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 14-JUN-1999; 99US-0138847P.
XX 16-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.

PR 25-AUG-1999;	99US-0150566P.	KW	Protein identification; signal transduction pathway; metabolic pathway;
PR 26-AUG-1999;	99US-0150884P.	KW	hybridisation assay; genetic mapping; gene expression control; promoter;
PR 27-AUG-1999;	99US-0151065P.	XX	termination sequence.
PR 27-AUG-1999;	99US-0151066P.	XX	
PR 27-AUG-1999;	99US-0151080P.	OS	Arabidopsis thaliana.
PR 30-AUG-1999;	99US-0151303P.	XX	
PR 31-AUG-1999;	99US-0151438P.	PN	EP1033405-A2.
PR 01-SEP-1999;	99US-0151930P.	XX	
PR 07-SEP-1999;	99US-0152363P.	PD	06-SEP-2000.
PR 10-SEP-1999;	99US-0153070P.	XX	
PR 13-SEP-1999;	99US-0153758P.	PF	
PR 15-SEP-1999;	99US-0154018P.	XX	
PR 16-SEP-1999;	99US-0154039P.	PR	25-FEB-2000; 2000EP-00301439.
PR 20-SEP-1999;	99US-0154779P.	PR	25-FEB-1999; 99US-0121825P.
PR 22-SEP-1999;	99US-0155139P.	PR	09-MAR-1999; 99US-0123180P.
PR 23-SEP-1999;	99US-0155486P.	PR	05-MAR-1999; 99US-0123548P.
PR 24-SEP-1999;	99US-0155659P.	PR	23-MAR-1999; 99US-0125788P.
PR 28-SEP-1999;	99US-0156458P.	PR	25-MAR-1999; 99US-0126264P.
PR 29-SEP-1999;	99US-0156596P.	PR	29-MAR-1999; 99US-0126785P.
PR 04-OCT-1999;	99US-0157117P.	PR	01-APR-1999; 99US-0127462P.
PR 05-OCT-1999;	99US-0157753P.	PR	06-APR-1999; 99US-0128234P.
PR 06-OCT-1999;	99US-0157865P.	PR	08-APR-1999; 99US-0128714P.
PR 07-OCT-1999;	99US-0158029P.	PR	16-APR-1999; 99US-0129845P.
PR 08-OCT-1999;	99US-0158232P.	PR	19-APR-1999; 99US-0130077P.
PR 12-OCT-1999;	99US-0158369P.	PR	21-APR-1999; 99US-0130449P.
PR 13-OCT-1999;	99US-0159293P.	PR	23-APR-1999; 99US-0130510P.
PR 13-OCT-1999;	99US-0159294P.	PR	23-APR-1999; 99US-0130891P.
PR 13-OCT-1999;	99US-0159295P.	PR	28-APR-1999; 99US-0131449P.
PR 14-OCT-1999;	99US-0159329P.	PR	30-APR-1999; 99US-0132048P.
PR 14-OCT-1999;	99US-0159330P.	PR	30-APR-1999; 99US-0132407P.
PR 14-OCT-1999;	99US-0159331P.	PR	04-MAY-1999; 99US-0132484P.
PR 14-OCT-1999;	99US-0159637P.	PR	05-MAY-1999; 99US-0132485P.
PR 14-OCT-1999;	99US-0159638P.	PR	06-MAY-1999; 99US-0132486P.
PR 18-OCT-1999;	99US-0159584P.	PR	07-MAY-1999; 99US-0132487P.
PR 21-OCT-1999;	99US-0160741P.	PR	07-MAY-1999; 99US-0132863P.
PR 21-OCT-1999;	99US-0160767P.	PR	11-MAY-1999; 99US-0134256P.
PR 21-OCT-1999;	99US-0160768P.	PR	14-MAY-1999; 99US-0134218P.
PR 21-OCT-1999;	99US-0160770P.	PR	14-MAY-1999; 99US-0134219P.
PR 21-OCT-1999;	99US-0160814P.	PR	14-MAY-1999; 99US-0134221P.
PR 21-OCT-1999;	99US-0160815P.	PR	18-MAY-1999; 99US-0134370P.
PR 22-OCT-1999;	99US-0160980P.	PR	18-MAY-1999; 99US-0134768P.
PR 22-OCT-1999;	99US-0160981P.	PR	19-MAY-1999; 99US-0134941P.
PR 22-OCT-1999;	99US-0160989P.	PR	20-MAY-1999; 99US-0135124P.
PR 25-OCT-1999;	99US-0161404P.	PR	21-MAY-1999; 99US-0135353P.
PR 25-OCT-1999;	99US-0161405P.	PR	21-MAY-1999; 99US-0135629P.
PR 25-OCT-1999;	99US-0161406P.	PR	25-MAY-1999; 99US-0136021P.
PR 26-OCT-1999;	99US-0161359P.	PR	27-MAY-1999; 99US-0136392P.
PR 26-OCT-1999;	99US-0161360P.	PR	28-MAY-1999; 99US-0136782P.
PR 26-OCT-1999;	99US-0161361P.	PR	01-JUN-1999; 99US-0137222P.
PR 28-OCT-1999;	99US-0161920P.	PR	03-JUN-1999; 99US-0137528P.
PR 28-OCT-1999;	99US-0161992P.	PR	04-JUN-1999; 99US-0137502P.
PR 28-OCT-1999;	99US-0161993P.	PR	07-JUN-1999; 99US-0137724P.
PR 29-OCT-1999;	99US-0162142P.	PR	08-JUN-1999; 99US-0138094P.
PR 29-OCT-1999;	99US-0162143P.	PR	10-JUN-1999; 99US-0138540P.
PR 29-OCT-1999;	99US-0162144P.	PR	10-JUN-1999; 99US-0138847P.
Query Match	100.0%; Score 35; DB 3; Length 115;	PR	14-JUN-1999; 99US-0139119P.
Best Local Similarity	100.0%; Pred. No. 4.2e+02;	PR	16-JUN-1999; 99US-0139452P.
Matches 5; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	PR	16-JUN-1999; 99US-0139453P.
QY	1 PPPHG 5	PR	17-JUN-1999; 99US-0139492P.
DB	79 PPPHG 83	PR	17-JUN-1999; 99US-0139493P.
		PR	18-JUN-1999; 99US-0139454P.
		PR	18-JUN-1999; 99US-0139455P.
RESULT 39		PR	18-JUN-1999; 99US-0139456P.
AAG16967		PR	18-JUN-1999; 99US-0139457P.
ID AAG16967 standard; protein; 116 AA.		PR	18-JUN-1999; 99US-0139458P.
XX AC		PR	18-JUN-1999; 99US-0139459P.
XX DT		PR	18-JUN-1999; 99US-0139460P.
XX DT		PR	18-JUN-1999; 99US-0139461P.
XX DT		PR	18-JUN-1999; 99US-0139462P.
XX DT		PR	18-JUN-1999; 99US-0139463P.
DE DT		PR	18-JUN-1999; 99US-0139750P.
XX DT		PR	18-JUN-1999; 99US-0139763P.
XX DT		PR	21-JUN-1999; 99US-0139817P.
XX DT		PR	21-JUN-1999; 99US-0139899P.
XX DT		PR	22-JUN-1999; 99US-0140353P.

PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR N-PSDB; AAI81904.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 20; SEQ ID NO 15865; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 118 AA;

Query Match 100.0%; Score 35; DB 4; Length 118;
Best Local Similarity 100.0%; Pred. No. 4.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
 |||||
Db 18 PPPHG 22

Search completed: June 15, 2005, 14:15:52
Job time : 92 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:06:20 ; Search time 23.5 Seconds
(without alignments)
15.883 Million cell updates/sec

Title: US-10-074-225A-10

Perfect score: 35

Sequence: 1 PPPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: /cgn2_6/ptodata/1/iaa/5A COMB pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	52	1	US-07-745-206A-28
2	35	100.0	52	2	Sequence 28, Appl
3	35	100.0	98	4	US-09-489-039A-10032
4	35	100.0	103	4	Sequence 10032, A
5	35	100.0	135	4	US-09-489-039A-12696
6	35	100.0	135	4	Sequence 12696, A
7	35	100.0	174	2	US-09-270-767-32232
8	35	100.0	174	2	Sequence 32232, A
9	35	100.0	174	3	US-08-683-262B-47
10	35	100.0	202	4	Sequence 47, Appl
11	35	100.0	240	5	US-09-361-707-47
12	35	100.0	267	4	Sequence 1165, Ap
13	35	100.0	321	4	PCT-US93-00227-4
14	35	100.0	332	4	Sequence 4, Appl1
15	35	100.0	332	4	US-09-270-767-43608
16	35	100.0	335	4	Sequence 43608, A
17	35	100.0	335	4	US-09-252-991A-32545
18	35	100.0	335	4	Sequence 32545, A
19	35	100.0	335	4	US-09-270-767-45004
20	35	100.0	335	4	Sequence 45004, A
21	35	100.0	354	4	US-08-894-139-4
22	35	100.0	354	4	Sequence 4, Appl1
23	35	100.0	434	4	US-09-949-016-11268
24	35	100.0	439	4	Sequence 11268, A
25	35	100.0	479	2	US-09-949-016-6701
26	35	100.0	491	2	Sequence 6701, Ap
27	35	100.0	523	1	US-09-248-796A-18484
28	35	100.0	523	1	Sequence 18484, A
29	35	100.0	523	1	US-08-149-097D-38
30	35	100.0	523	1	Sequence 38, Appl
31	35	100.0	523	1	US-10-029-180-106
32	35	100.0	523	1	Sequence 106, App
33	35	100.0	524	1	US-08-455-543A-42
34	35	100.0	524	1	Sequence 42, Appl
35	35	100.0	524	1	US-08-223-305C-42
36	35	100.0	524	1	Sequence 6, Appl
37	35	100.0	524	6	US-08-336-257A-6
38	35	100.0	524	6	Sequence 6, Appl
39	35	100.0	524	6	Patent No. 5386025
40	35	100.0	524	6	Sequence 9, Appl
41	35	100.0	554	3	US-08-895-590-9
42	35	100.0	554	3	Sequence 9, Appl
43	35	100.0	702	4	US-09-917-254-94
44	35	100.0	702	4	Sequence 94, Appl
45	35	100.0	1088	3	US-09-082-059-2
46	35	100.0	1088	3	Sequence 2, Appl1

28 35 100.0 1250 3 US-08-938-291A-9 Sequence 9, Appli
29 35 100.0 1250 4 US-09-589-619-9 Sequence 9, Appli
30 35 100.0 1495 4 US-09-697-898-3 Sequence 3, Appli
31 35 100.0 1512 4 US-09-697-898-2 Sequence 2, Appli
32 35 100.0 4377 4 US-09-949-016-6978 Sequence 6978, Ap
33 29 82.9 10 3 US-09-460-384-22 Sequence 22, Appl
34 29 82.9 12 2 US-08-556-597-138 Sequence 138, App
35 29 82.9 12 4 US-10-118-575A-17 Sequence 17, Appl
36 29 82.9 15 3 US-08-602-999A-382 Sequence 382, App
37 29 82.9 15 4 US-09-500-124-382 Sequence 382, App
38 29 82.9 16 6 5378805-5 Patent No. 5378805
39 29 82.9 16 6 5378805-5 Patent No. 5378805
40 29 82.9 17 2 US-08-982-597A-16 Sequence 16, Appl
41 29 82.9 17 3 US-09-136-218-16 Sequence 16, Appl
42 29 82.9 17 3 US-08-602-999A-416 Sequence 416, App
43 29 82.9 17 4 US-09-500-124-416 Sequence 416, App
44 29 82.9 18 3 US-08-602-999A-303 Sequence 303, App
45 29 82.9 18 3 US-08-602-999A-371 Sequence 371, App

ALIGNMENTS

RESULT 1

US-07-745-206A-28
; Sequence 28, Application US/07745206A
; Patent No. 5429921
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: McQue, Ann
; APPLICANT: Feldman, Daniel
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fitch, Even, Tabin & Flannery
; STREET: 135 S. LaSalle
; CITY: Chicago
; STATE: Illinois
; COUNTRY: U.S.A.
; ZIP: 60603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/745,206A
; FILING DATE: 19910815
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Feder, Scott B
; REFERENCE/DOCKET NUMBER: 51504
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 312-372-7842
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-07-745-206A-28

Query Match 100.0%; Score 35; DB 1; Length 52;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

Db 40 PPPHG 44

```
RESULT 2
; US-08-311-363-28
; Sequence 28, Application US/08311363
; Patent No. 5876958
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: Human Calcium Channel Compositions and
; TITLE OF INVENTION: Methods
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESS: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/311,363
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 6362-51506
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 52 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-311-363-28

Query Match 100.0%; Score 35; DB 2; Length 52;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 40 PPPHG 44

RESULT 3
US-09-489-039A-10032
; Sequence 10032, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 10032
; LENGTH: 98

Query Match 100.0%; Score 35; DB 4; Length 98;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 28 PPPHG 32

RESULT 4
US-09-489-039A-12696
; Sequence 12696, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 12696
; LENGTH: 103
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-12696

Query Match 100.0%; Score 35; DB 4; Length 103;
Best Local Similarity 100.0%; Pred. No. 50;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 10 PPPHG 14

RESULT 5
US-09-270-767-32232
; Sequence 32232, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 32232
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; US-09-270-767-32232

Query Match 100.0%; Score 35; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 36 PPPHG 40

RESULT 6
US-09-270-767-47449
; Sequence 47449, Application US/09270767
; Patent No. 6703491
```

```
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 47449
; LENGTH: 135
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
US-09-270-767-47449

Query Match          100.0%; Score 35; DB 4; Length 135;
Best Local Similarity 100.0%; Pred. No. 64;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      36 PPHG 40

RESULT 7
US-08-683-262B-47
; Sequence 47, Application US/08683262B
; Patent No. 5929220
; GENERAL INFORMATION:
; APPLICANT: Shuping Tong et al.
; TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
; NUMBER OF SEQUENCES: 75
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/683,262B
; FILING DATE: 18-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 31,819
; REFERENCE/DOCKET NUMBER: 00786/287002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-683-262B-47

Query Match          100.0%; Score 35; DB 2; Length 174;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      68 PPHG 72

RESULT 8
US-09-640-211A-1165
; Sequence 1165, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1165
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-640-211A-1165
```

```
US-09-361-707-47
; Sequence 47, Application US/09361707
; Patent No. 6258937
; GENERAL INFORMATION:
; APPLICANT: Tong, Shuping
; APPLICANT: Li, Jisu
; APPLICANT: Wande, Jack R.
; TITLE OF INVENTION: HEPA DNA VIRUS RECEPTOR
; NUMBER OF SEQUENCES: 117
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/361,707
; FILING DATE: 27-Jul-1999
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/683,262
; FILING DATE: 18-JUL-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Creason, Gary L.
; REGISTRATION NUMBER: 34,310
; REFERENCE/DOCKET NUMBER: 00786/287003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 47:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 174 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:
US-09-361-707-47

Query Match          100.0%; Score 35; DB 3; Length 174;
Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      68 PPHG 72

RESULT 9
US-09-640-211A-1165
; Sequence 1165, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1165
; LENGTH: 202
; TYPE: PRT
; ORGANISM: Pinus radiata
US-09-640-211A-1165
```

```
Query Match      100.0%; Score 35; DB 4; Length 202;
Best Local Similarity 100.0%; Pred. No. 93;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPHG 5
Db      196 PPHG 200

RESULT 10
PCT-US93-00227-4
; Sequence 4, Application PC/TUS9300227
; GENERAL INFORMATION:
; APPLICANT: Furneaux, Henry M.
; APPLICANT: Pogner, Jerome B.
; TITLE OF INVENTION: ANTIGEN RECOGNIZED BY PATIENTS WITH
; TITLE OF INVENTION: ANTIBODY ASSOCIATED LAMBERT-EATON MYASTHENIC SYNDROME
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: PCT/US93/00227
; APPLICATION NUMBER: PCT/US93/00227
; FILING DATE: 19930111
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 39874
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-977-9550
; TELEFAX: 212-664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 240 amino acids
; TYPE: AMINO ACID
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: internal
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..240
PCT-US93-00227-4

Query Match      100.0%; Score 35; DB 5; Length 240;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPHG 5
Db      222 PPHG 226

RESULT 11
US-09-270-767-43608
; Sequence 43608, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 43608
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: Xaa means any amino acid
US-09-270-767-43608

Query Match      100.0%; Score 35; DB 4; Length 267;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPHG 5
Db      224 PPHG 228

RESULT 12
US-09-252-991A-32545
; Sequence 32545, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32545
; LENGTH: 321
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32545

Query Match      100.0%; Score 35; DB 4; Length 321;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1 PPHG 5
Db      90 PPHG 94

RESULT 13
US-09-270-767-45004
; Sequence 45004, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 45004
; LENGTH: 332
; TYPE: PRT
```

```
; ORGANISM: Drosophila melanogaster
US-09-270-767-45004

Query Match      100.0%; Score 35; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      184 PPHG 188

RESULT 14
US-08-894-139-4
; Sequence 4, Application US/08894139
; Patent No. 6448376
; GENERAL INFORMATION:
; APPLICANT: LA THANGUE, NICHOLAS B.
; APPLICANT: BERNARDS, RENE
; APPLICANT: HIJMAN, ELEANORE M.
; TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 NORTH GLEBE ROAD
; CITY: ARLINGTON
; STATE: VIRGINIA
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/894,139
; FILING DATE: 13-AUG-1997
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J.
; REGISTRATION NUMBER: 32,955
; REFERENCE/DOCKET NUMBER: 620-22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 335 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-894-139-4

Query Match      100.0%; Score 35; DB 4; Length 335;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      20 PPHG 24

RESULT 15
US-09-949-016-11268
; Sequence 11268, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11268
; LENGTH: 354
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-11268

Query Match      100.0%; Score 35; DB 4; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      308 PPHG 312

RESULT 16
US-09-949-016-6701
; Sequence 6701, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6701
; LENGTH: 434
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6701

Query Match      100.0%; Score 35; DB 4; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      29 PPHG 33

RESULT 17
US-09-248-796A-18484
; Sequence 18484, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS
; FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13
; NUMBER OF SEQ ID NOS: 28208
```

```
/ SEQ ID NO 18484
/ LENGTH: 439
/ TYPE: PRT
/ ORGANISM: Candida albicans
US-09-248-796A-18484

Query Match      100.0%; Score 35; DB 4; Length 439;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      343 PPPHG 347

RESULT 18
US-08-149-097D-38
/ Sequence 38, Application US/08149097D
/ Patent No. 5874236
/ GENERAL INFORMATION:
/ APPLICANT: Harpold, Michael
/ APPLICANT: Ellis, Steven
/ APPLICANT: Williams, Mark
/ APPLICANT: Feldman, Daniel
/ APPLICANT: McCue, Ann
/ APPLICANT: Brenner, Robert
/ TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
/ METHODS
/ NUMBER OF SEQUENCES: 40
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Brown, Martin, Haller & McClain
/ STREET: 1660 Union Street
/ CITY: San Diego
/ STATE: California
/ COUNTRY: USA
/ ZIP: 92101-2926
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/149,097D
/ FILING DATE: 05-NOV-1993
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/105,536
/ FILING DATE: 11-AUG-1993
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/US92/06903
/ FILING DATE: 14-AUG-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/914,231
/ FILING DATE: 13-JUL-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/868,354
/ FILING DATE: 10-APR-1992
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/745,206
/ FILING DATE: 15-AUG-1991
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/620,250
/ FILING DATE: 30-NOV-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/482,384
/ FILING DATE: 20-FEB-1990
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US 07/603,751
/ FILING DATE: 04-APR-1989
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: WO PCT/US89/01408
/ FILING DATE: 04-APR-1989
/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: US 07/176,899
/ FILING DATE: 04-APR-1988
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Seidman, Stephanie L.
/ REGISTRATION NUMBER: 33,779
/ REFERENCE/DOCKET NUMBER: 6362-55038
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (619) 238-0999
/ TELEFAX: (619) 238-0062
/ INFORMATION FOR SEQ ID NO: 38:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 479 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ HYPOTHETICAL: NO
/ ANTI-SENSE: NO
/ FRAGMENT TYPE: internal
/ ORIGINAL SOURCE:
/ FEATURE:
/ OTHER INFORMATION: /standard_name= "Betal-1"
US-08-149-097D-38

Query Match      100.0%; Score 35; DB 2; Length 479;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      249 PPPHG 253

RESULT 19
US-10-029-180-106
/ Sequence 106, Application US/10029180
/ Patent No. 6806082
/ GENERAL INFORMATION:
/ APPLICANT: Cali, Brian M.
/ APPLICANT: Holtzman, Doug
/ APPLICANT: Madden, Kevin T.
/ APPLICANT: Milna, G. Todd
/ APPLICANT: Sherman, Amir
/ APPLICANT: Silva, Jeffry C.
/ APPLICANT: Trueheart, Josh
/ APPLICANT: Zhang, Lixin
/ TITLE OF INVENTION: No. 6806082el Regulators of Fungal Gene Expression
/ FILE REFERENCE: MIC-004
/ CURRENT APPLICATION NUMBER: US/10/029,180
/ CURRENT FILING DATE: 2001-12-22
/ PRIOR APPLICATION NUMBER: US 60/257,431
/ PRIOR FILING DATE: 2000-12-22
/ NUMBER OF SEQ ID NOS: 138
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 106
/ LENGTH: 491
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: fungal gene
US-10-029-180-106

Query Match      100.0%; Score 35; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      278 PPPHG 282

RESULT 20
US-08-455-543A-42
```

; Sequence 42, Application US/08455543A
; Patent No. 5792846
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/455,543A
; FILING DATE: May 31, 1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/223,305
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-455-543A-42

Query Match 100.0%; Score 35; DB 1; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5

Db 249 PPHG 253
RESULT 21
US-08-223-305C-42
; Sequence 42, Application US/08223305C
; Patent No. 5851824
; GENERAL INFORMATION:
; APPLICANT: Harpold, Michael
; APPLICANT: Ellis, Steven
; APPLICANT: Williams, Mark
; APPLICANT: Feldman, Daniel
; APPLICANT: McCue, Ann
; APPLICANT: Brenner, Robert
; TITLE OF INVENTION: HUMAN CALCIUM CHANNEL COMPOSITIONS AND
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: California
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/223,305C
; FILING DATE: April 4, 1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/868,354
; FILING DATE: April 10, 1992
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/745,206
; FILING DATE: 15-AUG-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/620,250
; FILING DATE: 30-NOV-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/482,384
; FILING DATE: 20-FEB-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/603,751
; FILING DATE: 04-APR-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/US89/01408
; FILING DATE: 04-APR-1989
; APPLICATION NUMBER: US 07/176,899
; FILING DATE: 04-APR-1988
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 52516 (P519739)
; TELEPHONE: (619)238-0999
; TELEFAX: (619)238-0062
; INFORMATION FOR SEQ ID NO: 42:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
US-08-223-305C-42

Query Match 100.0%; Score 35; DB 2; Length 523;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
|||||
Db 249 PPHG 253

RESULT 22

US-08-336-257A-6
; Sequence 6, Application US/08336257A
; Patent No. 5726035
; GENERAL INFORMATION:
; APPLICANT: Jay, Scott D
; APPLICANT: Ellis, Steven B.
; APPLICANT: Harpold, Michael M.
; APPLICANT: Campbell, Kevin P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Brown, Martin, Haller & McClain
; STREET: 1660 Union Street
; CITY: San Diego
; STATE: CA
; COUNTRY: USA
; ZIP: 92101-2926
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,257A
; FILING DATE: 07-NOV-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Seidman, Stephanie L.
; REGISTRATION NUMBER: 33,779
; REFERENCE/DOCKET NUMBER: 54898
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 238-0999
; TELEFAX: (619) 238-0062
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 524 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
US-08-336-257A-6

Query Match 100.0%; Score 35; DB 1; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
|||||
Db 249 PPHG 253

RESULT 23

5386025-4
; Patent No. 5386025
; APPLICANT: JAY, SCOTT D.; ELLIS, STEVEN B.; HARPOLD, MICHAEL
; M.; CAMPBELL, KEVIN P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,384
; FILING DATE: 20-FEB-1990
; SEQ ID NO: 4:
; LENGTH: 524

5386025-4

Query Match 100.0%; Score 35; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
|||||
Db 249 PPHG 253

RESULT 24

5386025-4
; Patent No. 5386025
; APPLICANT: JAY, SCOTT D.; ELLIS, STEVEN B.; HARPOLD, MICHAEL
; M.; CAMPBELL, KEVIN P.
; TITLE OF INVENTION: CALCIUM CHANNEL COMPOSITIONS AND METHODS
; NUMBER OF SEQUENCES: 9
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/482,384
; FILING DATE: 20-FEB-1990
; SEQ ID NO: 4:
; LENGTH: 524
5386025-4

Query Match 100.0%; Score 35; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
|||||
Db 249 PPHG 253

RESULT 25

US-08-895-590-9
; Sequence 9, Application US/08895590
; Patent No. 6207410
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,590
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,888
; FILING DATE: 19-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-263
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:

```
; LENGTH: 554 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-895-590-9
Query Match 100.0%; Score 35; DB 3; Length 554;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 250 PPHG 254

RESULT 26
US-09-917-254-94
; Sequence 94, Application US/09917254
; Patent No. 6703204
; GENERAL INFORMATION:
; APPLICANT: Mutter, George
; APPLICANT: Baak, Jan
; TITLE OF INVENTION: Prognostic Classification of Breast Cancer
; FILE REFERENCE: B0801/7224(JRV)
; CURRENT APPLICATION NUMBER: US/09/917,254
; CURRENT FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/222,093
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 102
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 94
; LENGTH: 702
; TYPE: PRT
; ORGANISM: Homo Sapiens
US-09-917-254-94

Query Match 100.0%; Score 35; DB 4; Length 702;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 228 PPHG 232

RESULT 27
US-09-082-059-2
; Sequence 2, Application US/09082059A
; Patent No. 6225086
; GENERAL INFORMATION:
; APPLICANT: Morrow, Jon S.
; APPLICANT: Devarajan, Prasad
; TITLE OF INVENTION: No. 6225086el Ankyrin Proteins and a Method for Their Identification
; FILE REFERENCE: 44574-5002-US
; CURRENT APPLICATION NUMBER: US/09/082,059A
; CURRENT FILING DATE: 1998-05-21
; EARLIER APPLICATION NUMBER: 60/047356
; EARLIER FILING DATE: 1997-05-21
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1088
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-082-059-2

Query Match 100.0%; Score 35; DB 3; Length 1088;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 228 PPHG 232

RESULT 28
US-08-938-291A-9
; Sequence 9, Application US/08938291A
; Patent No. 6117673
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDBG PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071-2066
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: FastSeq
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/938,291A
; FILING DATE: September 26, 1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/027,337
; FILING DATE: October 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Warburg, Richard J.
; REGISTRATION NUMBER: 32,327
; REFERENCE/DOCKET NUMBER: 228/172
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (213) 489-1600
; TELEFAX: (213) 955-0440
; TELEX: 67-3510
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1250 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-938-291A-9

Query Match 100.0%; Score 35; DB 3; Length 1250;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 435 PPHG 439

RESULT 29
US-09-589-619-9
; Sequence 9, Application US/09589619
; Patent No. 6576442
; GENERAL INFORMATION:
; APPLICANT: Lev, Sima
; APPLICANT: Plowman, Gregory D.
; APPLICANT: Schlessinger, Joseph
; TITLE OF INVENTION: RDBG PROTEINS AND RELATED
; TITLE OF INVENTION: PRODUCTS AND METHODS
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
```

ADDRESSEE: Lyon & Lyon
STREET: 633 West Fifth Street
Suite 4700
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90071-2066
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
COMPUTER: IBM Compatible
OPERATING SYSTEM: IBM P.C. DOS 5.0
SOFTWARE: FastSeq
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/589,619
FILING DATE: 07-Jun-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/938,291
FILING DATE: September 26, 1997
APPLICATION NUMBER: 60/027,337
FILING DATE: October 11, 1996
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 228/172
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 1250 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 9:

US-09-589-619-9

Query Match 100.0%; Score 35; DB 4; Length 1250;
Best Local Similarity 100.0%; Pred. No. 5.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
DB 435 PPHG 439

RESULT 30
US-09-697-898-3
Sequence 3, Application US/09697898
Patent No. 6818427
GENERAL INFORMATION:
APPLICANT: Liao, Sha-Mei
APPLICANT: Palombella, Vito J.
TITLE OF INVENTION: MEK1 Molecules and Uses Thereof
FILE REFERENCE: 103576.144
CURRENT APPLICATION NUMBER: US/09/697,898
CURRENT FILING DATE: 2000-10-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.0
SEQ ID NO 3
LENGTH: 1495
TYPE: PRT
ORGANISM: partial human MEK1 protein
US-09-697-898-3

Query Match 100.0%; Score 35; DB 4; Length 1495;
Best Local Similarity 100.0%; Pred. No. 6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5

DB 92 PPHG 96
RESULT 31
US-09-697-898-2
Sequence 2, Application US/09697898
Patent No. 6818427
GENERAL INFORMATION:
APPLICANT: Liao, Sha-Mei
APPLICANT: Palombella, Vito J.
TITLE OF INVENTION: MEK1 Molecules and Uses Thereof
FILE REFERENCE: 103576.144
CURRENT APPLICATION NUMBER: US/09/697,898
CURRENT FILING DATE: 2000-10-27
NUMBER OF SEQ ID NOS: 28
SOFTWARE: PatentIn version 3.0
SEQ ID NO 2
LENGTH: 1512
TYPE: PRT
ORGANISM: Human
US-09-697-898-2

Query Match 100.0%; Score 35; DB 4; Length 1512;
Best Local Similarity 100.0%; Pred. No. 6.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
DB 111 PPHG 115

RESULT 32
US-09-949-016-6978
Sequence 6978, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6978
LENGTH: 4377
TYPE: PRT
ORGANISM: Human
US-09-949-016-6978

Query Match 100.0%; Score 35; DB 4; Length 4377;
Best Local Similarity 100.0%; Pred. No. 1.6e+03;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
DB 1033 PPHG 1037

RESULT 33
US-09-460-384-22
Sequence 22, Application US/09460384
Patent No. 6337316
GENERAL INFORMATION:
APPLICANT: EL TAYAR, Nabil
BLECHNER, Steven

```
; JAMESON, Brad
; TEPPER, Mark
; TITLE OF INVENTION: CD28/CTLA-4 INHIBITING PEPTIDOMIMETICS,
; PHARMACEUTICAL COMPOSITIONS THEREOF, AND METHOD OF USING
; SAME
; NUMBER OF SEQUENCES: 37
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BROWDY AND NEIMARK, P.L.L.C.
; STREET: 624 Ninth Street N.W., Ste. 300
; CITY: Washington
; STATE: D.C.
; COUNTRY: U.S.A.
; ZIP: 20001
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/460,384
; FILING DATE: 13-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US98/12312
; FILING DATE: 11-JUN-1998
; APPLICATION NUMBER: US 60/049,470
; FILING DATE: 12-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: YUN, Allen C
; REGISTRATION NUMBER: 37,971
; REFERENCE/DOCKET NUMBER: EL TAYAR=1A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 628-5197
; TELEFAX: (202) 737-3528
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 10 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; SEQUENCE DESCRIPTION: SEQ ID NO: 22:
US-09-460-384-22
Query Match 82.9%; Score 29; DB 3; Length 10;
Best Local Similarity 80.0%; Pred. No. 53;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 4 PPPYG 8

RESULT 34
US-08-556-597-138
; Sequence 138, Application US/08556597
; Patent No. 5877155
; GENERAL INFORMATION:
; APPLICANT: Miller, Jonathan L.
; APPLICANT: Lyle, Vicki A.
; TITLE OF INVENTION: MIMOTOPES AND ANTI-MIMOTOPES OF
; TITLE OF INVENTION: HUMAN PLATELET GLYCOPROTEIN Ib/IX
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: USA
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
```

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/556,597
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/406,330
; FILING DATE: 17-MAR-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Timian, Susan J.
; REGISTRATION NUMBER: 34,103
; REFERENCE/DOCKET NUMBER: 20884/101
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1636
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 138:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-556-597-138
Query Match 82.9%; Score 29; DB 2; Length 12;
Best Local Similarity 80.0%; Pred. No. 63;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 6 PPPYG 10

RESULT 35
US-10-118-575A-17
; Sequence 17, Application US/10118575A
; Patent No. 6653443
; GENERAL INFORMATION:
; APPLICANT: ZHANG, Hui
; APPLICANT: PONERANZ, Roger
; APPLICANT: YANG, Bin
; TITLE OF INVENTION: Multimerization of HIV-1 VIF Protein as
; TITLE OF INVENTION: a Therapeutic Target
; FILE REFERENCE: 8321-82 PC
; CURRENT APPLICATION NUMBER: US/10/118,575A
; CURRENT FILING DATE: 2002-04-08
; PRIOR APPLICATION NUMBER: US 60/282,270
; PRIOR FILING DATE: 2001-04-06
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 12
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic peptide containing PXP motif
; US-10-118-575A-17
Query Match 82.9%; Score 29; DB 4; Length 12;
Best Local Similarity 100.0%; Pred. No. 63;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPH 4
DB 2 PPPH 5

RESULT 36
US-08-602-999A-382
; Sequence 382, Application US/08602999A
; Patent No. 6184205
; GENERAL INFORMATION:
```

```
/ APPLICANT: SPARKS, Andrew B.
/ APPLICANT: KAY, Brian K.
/ APPLICANT: THORN, Judith M.
/ APPLICANT: QUILLIAM, Lawrence A.
/ APPLICANT: DER, Channing J.
/ APPLICANT: FOWKES, Dana M.
/ APPLICANT: RIDER, James E.
/ TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
/ TITLE OF INVENTION: ISOLATING AND USING SAME
/ NUMBER OF SEQUENCES: 467
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/602,999A
/ FILING DATE: 16-FEB-1996
/ CLASSIFICATION: 435
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mirock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 1101-202
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 382:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: peptide
/ US-08-602-999A-382

Query Match      82.9%; Score 29; DB 3; Length 15;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPH 4
        ||||
Db      9 PPPH 12

RESULT 37
US-09-500-124-382
/ Sequence 382, Application US/09500124
/ Patent No. 6432920
/ GENERAL INFORMATION:
/ APPLICANT: SPARKS, Andrew B.
/ APPLICANT: KAY, Brian K.
/ APPLICANT: THORN, Judith M.
/ APPLICANT: QUILLIAM, Lawrence A.
/ APPLICANT: DER, Channing J.
/ APPLICANT: FOWKES, Dana M.
/ APPLICANT: RIDER, James E.
/ TITLE OF INVENTION: SH3 BINDING PEPTIDES AND METHODS OF
/ TITLE OF INVENTION: ISOLATING AND USING SAME
/ NUMBER OF SEQUENCES: 467
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: Pennie & Edmonds
/ STREET: 1155 Avenue of the Americas
/ CITY: New York
/ STATE: New York
/ COUNTRY: U.S.A.
/ ZIP: 10036-2711
```

```
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/500,124
/ FILING DATE:
/ CLASSIFICATION:
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: 08/602,999
/ FILING DATE: 16-FEB-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Mirock, S. Leslie
/ REGISTRATION NUMBER: 18,872
/ REFERENCE/DOCKET NUMBER: 1101-202
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (212) 790-9090
/ TELEFAX: (212) 869-9741/8864
/ TELEX: 66141 PENNIE
/ INFORMATION FOR SEQ ID NO: 382:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 15 amino acids
/ TYPE: amino acid
/ TOPOLOGY: unknown
/ MOLECULE TYPE: peptide
/ US-09-500-124-382

Query Match      82.9%; Score 29; DB 4; Length 15;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPH 4
        ||||
Db      9 PPPH 12

RESULT 38
5378805-5
/ Patent No. 5378805
/ APPLICANT: Lai, Renu B.
/ TITLE OF INVENTION: IMMUNOREACTIVE HTLV-I/II ENV AND POL PEPTIDES
/ NUMBER OF SEQUENCES: 5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/574,352
/ FILING DATE: 29-AUG-1990
/ SEQ ID NO:5:
/ LENGTH: 16
/ 5378805-5

Query Match      82.9%; Score 29; DB 6; Length 16;
Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPH 4
        ||||
Db      12 PPPH 15

RESULT 39
5378805-5
/ Patent No. 5378805
/ APPLICANT: Lai, Renu B.
/ TITLE OF INVENTION: IMMUNOREACTIVE HTLV-I/II ENV AND POL PEPTIDES
/ NUMBER OF SEQUENCES: 5
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/07/574,352
/ FILING DATE: 29-AUG-1990
/ SEQ ID NO:5:
/ LENGTH: 16
/ 5378805-5

Query Match      82.9%; Score 29; DB 6; Length 16;
```

Best Local Similarity 100.0%; Pred. No. 83;
Matches 4; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPH 4
Db 12 PPH 15

RESULT 40

US-08-982-597A-16
; Sequence 16, Application US/08982597A
; Patent No. 5932693
; GENERAL INFORMATION:
; APPLICANT: Santoro, Samuel A.
; APPLICANT: Staats, William D.
; TITLE OF INVENTION: Antithrombotic Peptides
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Scott J. Meyer
; STREET: 800 No. 5932693th Lindbergh Blvd.
; CITY: St. Louis
; STATE: MO
; COUNTRY: USA
; ZIP: 63167
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 5.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/982,597A
; FILING DATE:
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/032,542
; FILING DATE: 10-DEC-1996
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: Meyer, Scott J.
; REGISTRATION NUMBER: 25,275
; REFERENCE/DOCKET NUMBER: WU-3002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 314-694-3117
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 17 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-982-597A-16

Query Match 82.9%; Score 29; DB 2; Length 17;
Best Local Similarity 80.0%; Pred. No. 87;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 12 PPYG 16

Search completed: June 15, 2005, 14:24:21
Job time : 24.5 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:21:52 ; Search time 79.75 Seconds
(without alignments)
24.034 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1710399 seqs, 383334425 residues

Total number of hits satisfying chosen parameters: 1710399

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:**

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10E_PUBCOMB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US11A_PUBCOMB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 21: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 22: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	35	100.0	5	14	US-10-074-225A-10
2	35	100.0	40	16	US-10-425-115-307567
3	35	100.0	44	16	US-10-437-963-139202
4	35	100.0	48	16	US-10-425-115-340254
5	35	100.0	51	16	US-10-425-115-199641
6	35	100.0	53	16	US-10-425-115-242501
7	35	100.0	55	9	US-09-864-761-34953
8	35	100.0	55	15	US-10-424-599-170955
9	35	100.0	56	16	US-10-425-115-199279
10	35	100.0	58	16	US-10-437-963-152655
11	35	100.0	58	16	US-10-425-115-323782

12	35	100.0	59	16	US-10-425-115-261805
13	35	100.0	60	16	US-10-425-115-352770
14	35	100.0	61	9	US-09-864-761-46134
15	35	100.0	61	16	US-10-425-115-239349
16	35	100.0	62	16	US-10-425-115-344565
17	35	100.0	65	16	US-10-424-599-206662
18	35	100.0	65	16	US-10-425-115-316672
19	35	100.0	68	16	US-10-437-963-130521
20	35	100.0	70	16	US-10-437-963-150972
21	35	100.0	72	15	US-10-424-599-181807
22	35	100.0	72	16	US-10-425-115-185365
23	35	100.0	73	15	US-10-424-599-187344
24	35	100.0	75	16	US-10-767-701-51220
25	35	100.0	75	16	US-10-425-115-307219
26	35	100.0	76	16	US-10-425-115-258443
27	35	100.0	77	16	US-10-437-963-158099
28	35	100.0	79	16	US-10-437-963-132911
29	35	100.0	80	15	US-10-424-599-220866
30	35	100.0	80	16	US-10-425-115-249543
31	35	100.0	80	16	US-10-425-115-323576
32	35	100.0	81	16	US-10-437-963-109589
33	35	100.0	81	16	US-10-425-115-233659
34	35	100.0	82	16	US-10-425-115-271683
35	35	100.0	84	15	US-10-424-599-163493
36	35	100.0	84	15	US-10-424-599-275367
37	35	100.0	84	16	US-10-437-963-133341
38	35	100.0	86	16	US-10-425-115-204320
39	35	100.0	88	15	US-10-424-599-143090
40	35	100.0	88	15	US-10-424-599-260730
41	35	100.0	89	15	US-10-282-122A-62241
42	35	100.0	89	16	US-10-425-115-232815
43	35	100.0	90	16	US-10-437-963-123395
44	35	100.0	91	15	US-10-424-599-151231
45	35	100.0	91	15	US-10-424-599-252891

ALIGNMENTS

RESULT 1

US-10-074-225A-10
; Sequence 10, Application US/10074225A
; Publication No. US20030082740A1
; GENERAL INFORMATION:
; APPLICANT: DONATE, Fernando
; APPLICANT: PLUNKETT, Marian L
; APPLICANT: HARRIS, Scott
; APPLICANT: MAZAR, Andrew P
; TITLE OF INVENTION: HISTIDINE PROLINE RICH GLYCOPROTEIN (HPRG) AS AN ANTI-ANGIOGENIC I
; FILE REFERENCE: 38342-178463
; CURRENT FILING DATE: 2002-02-14
; PRIOR FILING DATE: 2001-02-14
; NUMBER OF SEQ ID NOS: 11
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 10
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic Peptide
US-10-074-225A-10

Query Match 100.0% Score 35; DB 14; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.5e+06; Indels 0; Gaps 0;
Matches 5; Conservative 0; Mismatches 0;

QY 1 PPHG 5

Db 1 PPHG 5

```
RESULT 2
US-10-425-115-307567
; Sequence 307567, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 307567
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_43570C.1.pap
US-10-425-115-307567

Query Match      100.0%; Score 35; DB 16; Length 40;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      14 PPPHG 18

RESULT 3
US-10-437-963-139202
; Sequence 139202, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 139202
; LENGTH: 44
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(44)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_40517C.1.pap
US-10-437-963-139202

Query Match      100.0%; Score 35; DB 16; Length 44;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      30 PPPHG 34

RESULT 4
US-10-425-115-340254
; Sequence 340254, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 340254
; LENGTH: 48
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_73481C.1.pap
US-10-425-115-340254

Query Match      100.0%; Score 35; DB 16; Length 48;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      43 PPPHG 47

RESULT 5
US-10-425-115-199641
; Sequence 199641, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199641
; LENGTH: 51
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113645C.1.pap
US-10-425-115-199641

Query Match      100.0%; Score 35; DB 16; Length 51;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
DB      31 PPPHG 35

RESULT 6
US-10-425-115-242501
; Sequence 242501, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 242501
; LENGTH: 53
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_152747C.1.pap
US-10-425-115-242501

Query Match 100.0%; Score 35; DB 16; Length 53;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 31 PPPHG 35

RESULT 7

US-09-864-761-34953
; Sequence 34953, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aecm1ca-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29

; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 34953
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC004900.2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 3.7
; OTHER INFORMATION: EXPRESSED IN BT474, SIGNAL = 4.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 7.1
; OTHER INFORMATION: EXPRESSED IN HBL100, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN HEAT, SIGNAL = 5.9
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 4.4
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 3.1
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 2.5
; OTHER INFORMATION: EST HUMAN HIT: AW953375.1, EVALUE 6.00e-29
; OTHER INFORMATION: SWISSPROT HIT: P16793, EVALUE 5.20e-01
US-09-864-761-34953

Query Match 100.0%; Score 35; DB 9; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 1 PPPHG 5

RESULT 8

US-10-424-599-170955
; Sequence 170955, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 170955
; LENGTH: 55
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_125388C.1.pap
US-10-424-599-170955

Query Match 100.0%; Score 35; DB 15; Length 55;
Best Local Similarity 100.0%; Pred. No. 2.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 33 PPPHG 37

RESULT 9

US-10-425-115-199279
; Sequence 199279, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 199279
; LENGTH: 56
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_113318C.1.pbp
US-10-425-115-199279

Query Match      100.0%; Score 35; DB 16; Length 56;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      1 PPHG 5

RESULT 10
US-10-437-963-152655
; Sequence 152655, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 152655
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_52685C.1.pbp
US-10-437-963-152655

Query Match      100.0%; Score 35; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      21 PPHG 25

RESULT 11
US-10-425-115-323782
; Sequence 323782, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 323782
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84894C.1.pbp
US-10-425-115-323782

Query Match      100.0%; Score 35; DB 16; Length 60;
```

```
; SEQ ID NO 323782
; LENGTH: 58
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_5835C.1.pbp
US-10-425-115-323782

Query Match      100.0%; Score 35; DB 16; Length 58;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      11 PPHG 15

RESULT 12
US-10-425-115-261805
; Sequence 261805, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 261805
; LENGTH: 59
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_170379C.1.pbp
US-10-425-115-261805

Query Match      100.0%; Score 35; DB 16; Length 59;
Best Local Similarity 100.0%; Pred. No. 2.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
DB      49 PPHG 53

RESULT 13
US-10-425-115-352770
; Sequence 352770, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 352770
; LENGTH: 60
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_84894C.1.pbp
US-10-425-115-352770

Query Match      100.0%; Score 35; DB 16; Length 60;
```

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 31 PPPHG 35

RESULT 14
US-09-864-761-46134
; Sequence 46134, Application US/09864761
; Patent No. US20020048763A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
; FILE REFERENCE: Aeomica-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 46134
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC011625-2
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 8.5
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1.6

US-09-864-761-46134

Query Match 100.0%; Score 35; DB 9; Length 61;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 18 PPPHG 22

RESULT 15
US-10-425-115-239349
; Sequence 239349, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 239349
; LENGTH: 61
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_149873C.1.pap

US-10-425-115-239349

Query Match 100.0%; Score 35; DB 16; Length 61;

Best Local Similarity 100.0%; Pred. No. 2.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 2 PPPHG 6

RESULT 16
US-10-425-115-344565
; Sequence 344565, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 344565
; LENGTH: 62
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_77405C.1.pap

US-10-425-115-344565

Query Match 100.0%; Score 35; DB 16; Length 62;

Best Local Similarity 100.0%; Pred. No. 2.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 27 PPPHG 31

RESULT 17

US-10-424-599-206662
; Sequence 206662, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 206662
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_28643C.1.pap
US-10-424-599-206662

Query Match 100.0%; Score 35; DB 15; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 55 PPPHG 59

RESULT 18
US-10-425-115-316672
; Sequence 316672, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 316672
; LENGTH: 65
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_51874C.1.pap
US-10-425-115-316672

Query Match 100.0%; Score 35; DB 16; Length 65;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 25 PPPHG 29

RESULT 19
US-10-437-963-130521
; Sequence 130521, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 130521
; LENGTH: 68
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_32676C.1.pap
US-10-437-963-130521

Query Match 100.0%; Score 35; DB 16; Length 68;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 48 PPPHG 52

RESULT 20
US-10-437-963-150972
; Sequence 150972, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 150972
; LENGTH: 70
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_51159C.1.pap
US-10-437-963-150972

Query Match 100.0%; Score 35; DB 16; Length 70;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 13 PPPHG 17

RESULT 21
US-10-424-599-181807
; Sequence 181807, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 181807
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_135185C.1.pap
US-10-424-599-181807

Query Match 100.0%; Score 35; DB 15; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 7 PPHG 11

RESULT 22

US-10-425-115-185365
; Sequence 185365, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 185365
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(72)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_100640C.1.pap
US-10-425-115-185365

Query Match 100.0%; Score 35; DB 16; Length 72;
Best Local Similarity 100.0%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 42 PPHG 46

RESULT 23

US-10-424-599-187344
; Sequence 187344, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 187344

; LENGTH: 73
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_140183C.1.pap
US-10-424-599-187344

Query Match 100.0%; Score 35; DB 15; Length 73;
Best Local Similarity 100.0%; Pred. No. 3.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 69 PPHG 73

RESULT 24

US-10-767-701-51220
; Sequence 51220, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 51220
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3481-060-Q6-K1-D12.pap
US-10-767-701-51220

Query Match 100.0%; Score 35; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 8 PPHG 12

RESULT 25

US-10-425-115-307219
; Sequence 307219, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 307219
; LENGTH: 75
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(75)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_43257C.1.pap
US-10-425-115-307219

Query Match 100.0%; Score 35; DB 16; Length 75;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
DB 5 PPPHG 9

RESULT 26
US-10-425-115-258243
; Sequence 258243, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 258243
; LENGTH: 76
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_167102C.1.pep
US-10-425-115-258243

Query Match 100.0%; Score 35; DB 16; Length 76;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
DB 29 PPPHG 33

RESULT 27
US-10-437-963-158099
; Sequence 158099, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 158099
; LENGTH: 77
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_57606C.1.pep
US-10-437-963-158099

Query Match 100.0%; Score 35; DB 16; Length 77;
Best Local Similarity 100.0%; Pred. No. 3.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
DB 16 PPPHG 20

RESULT 28
US-10-437-963-132911
; Sequence 132911, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 132911
; LENGTH: 79
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_34833C.1.pep
US-10-437-963-132911

Query Match 100.0%; Score 35; DB 16; Length 79;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
DB 37 PPPHG 41

RESULT 29
US-10-424-599-220866
; Sequence 220866, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 220866
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(80)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_41472C.1.pep
US-10-424-599-220866

Query Match 100.0%; Score 35; DB 15; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5

```
Db          44 PPHG 48
|||||
RESULT 30
US-10-425-115-249543
; Sequence 249543, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 109589
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_13731C.1.pap
US-10-437-963-109589

Query Match          100.0%; Score 35; DB 16; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      75 PPHG 79
|||||

RESULT 33
US-10-425-115-233659
; Sequence 233659, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 233659
; LENGTH: 81
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_144690C.1.pap
US-10-425-115-233659

Query Match          100.0%; Score 35; DB 16; Length 81;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      56 PPHG 60
|||||

RESULT 34
US-10-425-115-271683
; Sequence 271683, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
```

```
Db          44 PPHG 48
|||||
RESULT 30
US-10-425-115-249543
; Sequence 249543, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 249543
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_159168C.1.pap
US-10-425-115-249543

Query Match          100.0%; Score 35; DB 16; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      12 PPHG 16
|||||

RESULT 31
US-10-425-115-323576
; Sequence 323576, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 323576
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_58171C.1.pap
US-10-425-115-323576

Query Match          100.0%; Score 35; DB 16; Length 80;
Best Local Similarity 100.0%; Pred. No. 3.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPHG 5
Db      8 PPHG 12
|||||

RESULT 32
US-10-437-963-109589
; Sequence 109589, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
```

```
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 271683
; LENGTH: 82
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_179368C.1.pep
US-10-425-115-271683

Query Match      100.0%; Score 35; DB 16; Length 82;
Best Local Similarity 100.0%; Pred. No. 3.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      44 PPPHG 48

RESULT 35
US-10-424-599-163493
; Sequence 163493, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 163493
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_118653C.1.pep
US-10-424-599-163493

Query Match      100.0%; Score 35; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      38 PPPHG 42

RESULT 36
US-10-424-599-275367
; Sequence 275367, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
```

```
; SEQ ID NO 275367
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90677C.1.pep
US-10-424-599-275367

Query Match      100.0%; Score 35; DB 15; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      69 PPPHG 73

RESULT 37
US-10-437-963-133341
; Sequence 133341, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 133341
; LENGTH: 84
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(84)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_35220C.1.pep
US-10-437-963-133341

Query Match      100.0%; Score 35; DB 16; Length 84;
Best Local Similarity 100.0%; Pred. No. 3.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      7 PPPHG 11

RESULT 38
US-10-425-115-204320
; Sequence 204320, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 204320
```

```
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_117932C.1.pep
US-10-425-115-204320

Query Match      100.0%; Score 35; DB 16; Length 86;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      |||||
      82 PPPHG 86

RESULT 39
US-10-424-599-143090
; Sequence 143090, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 143090
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_100223C.1.pep
US-10-424-599-143090

Query Match      100.0%; Score 35; DB 15; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      |||||
      7 PPPHG 11

RESULT 40
US-10-424-599-260730
; Sequence 260730, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 260730
; LENGTH: 88
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_77461C.1.pep
US-10-424-599-260730

Query Match      100.0%; Score 35; DB 15; Length 88;
Best Local Similarity 100.0%; Pred. No. 3.8e+02;
```

```
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1 PPPHG 5
Db      |||||
      61 PPPHG 65
```

Search completed: June 15, 2005, 14:50:21
Job time : 81.75 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (C) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 14:04:57 ; Search time 17.5 Seconds
(without alignments)
27.491 Million cell updates/sec

Title: US-10-074-225A-10

Perfect score: 35

Sequence: 1 PPPHG 5

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: Pirl:*

2: Pirl:*

3: Pirl:*

4: Pirl:*

pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	100.0	111	2 C72714	hypothetical prote
2	35	100.0	147	2 S37485	gene msgl protein
3	35	100.0	168	2 C87235	conserved hypothet
4	35	100.0	169	2 D70906	hypothetical prote
5	35	100.0	170	2 A48013	proline-rich prote
6	35	100.0	173	2 T51469	glycine/proline-ri
7	35	100.0	173	2 T19341	hypothetical prote
8	35	100.0	175	2 T46665	hypothetical prote
9	35	100.0	176	2 A86441	hypothetical prote
10	35	100.0	206	1 PIRT3	acidic proline-ric
11	35	100.0	215	2 F72641	hypothetical prote
12	35	100.0	224	2 B87440	transcription regu
13	35	100.0	257	2 T10586	small nuclear ribo
14	35	100.0	278	1 EDBE63	immediate-early pr
15	35	100.0	301	2 T35805	hypothetical prote
16	35	100.0	304	2 A55852	membrane-associate
17	35	100.0	316	2 S16881	homeotic protein -
18	35	100.0	335	2 I48338	E2F-5 - mouse
19	35	100.0	366	2 T26449	hypothetical prote
20	35	100.0	368	2 G84769	hypothetical prote
21	35	100.0	391	2 B87723	protein R06A10.4 [
22	35	100.0	422	2 T07967	mandelonitrile ly
23	35	100.0	429	2 T32632	hypothetical prote
24	35	100.0	434	2 S19010	homeotic protein P
25	35	100.0	454	2 F81046	hypothetical prote
26	35	100.0	459	2 B69458	sulfate adenyllylr
27	35	100.0	465	2 G02738	FRAC-4 - human
28	35	100.0	475	2 A44461	voltage-dependent
29	35	100.0	486	2 T20482	hypothetical prote

30	35	100.0	509	2 JC6203	SP8 binding protei
31	35	100.0	516	2 T49422	RAU57-related prot
32	35	100.0	518	2 S61920	B-alpha pheromone-
33	35	100.0	522	2 I65767	L-type voltage-gat
34	35	100.0	524	2 A41347	calcium channel pr
35	35	100.0	529	2 T00677	hypothetical prote
36	35	100.0	597	2 G84825	probable CCH-type
37	35	100.0	614	2 E86194	hypothetical prote
38	35	100.0	627	2 S48424	hypothetical prote
39	35	100.0	629	2 C87048	probable ABC trans
40	35	100.0	669	2 A49585	Na+ channel protei
41	35	100.0	830	2 T36683	probable integral
42	35	100.0	960	2 A41638	chitin synthase [S
43	35	100.0	968	2 F70746	probable mmpL2 pro
44	35	100.0	968	2 T00353	hypothetical prote
45	35	100.0	1048	2 T31425	C-terminal domain-

ALIGNMENTS

RESULT 1

C72714

hypothetical protein APE1130 - Aeropyrum pernix (strain K1)

C:Species: Aeropyrum pernix

C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004

C:Accession: C72714

R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takah

awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; K

DNA Res. 6, 83-101, 1999

A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr

A:Reference number: A72450; MUID:99310339; PMID:10382966

A:Accession: C72714

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-111 <KAW>

A:Cross-references: UNIPROT:Q9YCY2; DDBJ:AP000060; NID:G5104188; PIDN:BAA80115.1; PID:d1

A:Experimental source: strain K1

C:Genetics:

A:Gene: APE1130

Query Match 100.0%; Score 35; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PPPHG 5

Db 82 PPPHG 86

RESULT 2

S37485

gene msgl protein - mouse

C:Species: Mus musculus (house mouse)

C>Date: 06-Jan-1995 #sequence_revision 06-Jan-1995 #text_change 09-Jul-2004

C:Accession: I48669; S37485

R:Tronik-Le Roux, D.; Senorale-Pose, M.; Rougeon, F.

Gene 142, 175-182, 1994

A:Title: Three novel SMR1-related cDNAs characterized in the submaxillary gland of mice

A:Reference number: I48669; MUID:94252564; PMID:8194749

A:Accession: I48669

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-147 <RES>

A:Cross-references: UNIPROT:Q61900; EMBL:X71629; NID:G406256; PIDN:CAA50636.1; PID:G4062

C:Genetics:

A:Gene: msgl

C:Superfamily: proline-rich peptide P-B

Query Match 100.0%; Score 35; DB 2; Length 147;

Best Local Similarity 100.0%; Pred. No. 38;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 39 PPPHG 43

RESULT 3

C87235 conserved hypothetical protein ML2605 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87235
R: Cole, S.T.; Egilmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Holt,
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Squ
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C87235
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-168 <STO>
A:Cross-references: UNIPROT:O06065; GB:AL450380; NID:g13093836; PIDN:CAC32137.1; GSPDB:C
C:Genetics:
A:Gene: ML2605
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0185

Query Match 100.0%; Score 35; DB 2; Length 168;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 131 PPPHG 135

RESULT 4

D70906 hypothetical protein Rv0185 - Mycobacterium tuberculosis (strain H37Rv)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70906
R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70906
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-169 <COL>
A:Cross-references: UNIPROT:O07429; GB:Z97050; GB:AL123456; NID:g3256008; PIDN:CAB09736.
A:Experimental source: strain H37Rv
C:Genetics:
A:Gene: Rv0185
C:Superfamily: Mycobacterium tuberculosis hypothetical protein Rv0185

Query Match 100.0%; Score 35; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 43;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 132 PPPHG 136

RESULT 5

A48013 proline-rich proteoglycan 1 precursor, parotid - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 02-Jun-1995 #sequence_revision 02-Jun-1995 #text_change 09-Jul-2004

C:Accession: A48013
R: Castle, A.M.; Castle, J.D.
J. Biol. Chem. 268, 20490-20496, 1993
A:Title: Novel secretory proline-rich proteoglycans from rat parotid. Cloning and charact
A:Reference number: A48013; MUID:93388626; PMID:8376404
A:Accession: A48013
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-170 <CAS>
A:Cross-references: UNIPROT:O07610; GB:L17317; NID:g310197; PIDN:AAA03073.1; PID:g310198
C:Superfamily: proline-rich protein
C:Keywords: extracellular protein; glycoprotein; tandem repeat
Query Match 100.0%; Score 35; DB 2; Length 170;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 80 PPPHG 84

RESULT 6

T51469 glycine/proline-rich protein - Arabidopsis thaliana
N:Alternate names: protein K10A8_130
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 09-Jul-2004
C:Accession: T51469
R: Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asanizu, E.; Kotani, H.; Tabata, S.; Mew
submitted to the Protein Sequence Database, August 2000
A:Reference number: Z25394
A:Accession: T51469
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-173 <SAT>
A:Cross-references: UNIPROT:Q9LF59; EMBL:AL391151
A:Experimental source: cultivar Columbia; BAC clone K10A8
C:Genetics:
A:Map position: 5
A:Introns: 97/1
A:Note: K10A8_130

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||||
Db 49 PPPHG 53

RESULT 7

T19341 hypothetical protein Cl6D6.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T19341
R: Gardner, A.
submitted to the EMBL Data Library, November 1996
A:Reference number: Z19111
A:Accession: T19341
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-173 <WILL>
A:Cross-references: UNIPROT:O62063; EMBL:Z81472; PIDN:CAB03889.1; GSPDB:GN000028; CESP:C1
A:Experimental source: clone Cl6D6
C:Genetics:
A:Gene: CESP:Cl6D6.1
A:Map position: X
A:Introns: 42/1

Query Match 100.0%; Score 35; DB 2; Length 173;

```

Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 72 PPHG 76

RESULT 8
T46665
Hypothetical protein [imported] - Rhodococcus corallinus (fragment)
C:Species: Rhodococcus corallinus
C>Date: 18-Feb-2000 #sequence_revision 18-Feb-2000 #text_change 09-Jul-2004
C:Accession: T46665
R:Shao, Z.Q.; Seffens, W.; Mulbry, W.; Behki, R.M.
J. Bacteriol. 177, 5748-5755, 1995
A:Title: Cloning and expression of the s-triazine hydrolase gene (trza) from Rhodococcus
erbicid atrazine.
A:Reference number: Z23125; MUID:96011356; PMID:7592318
A:Accession: T46665
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-175 <SHA>
A:Cross-references: UNIPROT:Q52724; EMBL:LI6534; NID:g294669; PIDN:AAA90930.1; PID:g2946
A:Experimental source: tissue lib NRRL 15444B

Query Match 100.0%; Score 35; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 129 PPHG 133

RESULT 9
A86441
Hypothetical protein F5M6.24 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: A86441
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: A86441
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-176 <StO>
A:Cross-references: UNIPROT:Q9C428; GB:AE005172; NID:g1136740; PIDN:AAG31321.1; GSPDB:G
C:Genetics:
A:Map position: 1

Query Match 100.0%; Score 35; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 50 PPHG 54

RESULT 10
P1RT3
acidic proline-rich protein precursor - rat
N:Alternate names: PRP

```

```

C:Species: Rattus norvegicus (Norway rat)
C>Date: 17-Mar-1987 #sequence_revision 17-Mar-1987 #text_change 09-Jul-2004
C:Accession: A03296
R:Ziemer, M.A.; Swain, W.F.; Rutter, W.J.; Clements, S.; Ann, D.K.; Carlson, D.M.
J. Biol. Chem. 259, 10475-10480, 1984
A:Title: Nucleotide sequence analysis of a proline-rich protein cDNA and peptide homologi
A:Reference number: A03296; MUID:84289443; PMID:6547951
A:Accession: A03296
A:Molecule type: mRNA
A:Residues: 1-206 <ZIE>
A:Cross-references: UNIPROT:P04474; GB:K02247; NID:g206395; PIDN:AAA41949.1; PID:g206396
C:Comment: This protein contains six 18- to 19-residue repeats.
C:Comment: This protein may protect teeth by binding to tannins.
C:Superfamily: proline-rich protein
C:Keywords: duplication; parotid gland; saliva; tandem repeat
F:14-206/Product: acidic proline-rich protein #status predicted <SIG>
F:80-189/Region: 18-residue repeats

Query Match 100.0%; Score 35; DB 1; Length 206;
Best Local Similarity 100.0%; Pred. No. 53;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 80 PPHG 84

RESULT 11
F72641
Hypothetical protein APE0566 - Aeropyrum pernix (strain K1)
C:Species: Aeropyrum pernix
C>Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
C:Accession: F72641
R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takaha
awa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Ka
DNA Res. 6, 83-101, 1999
A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyr
A:Reference number: A72450; MUID:99310339; PMID:10382966
A:Accession: F72641
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-215 <KAW>
A:Cross-references: UNIPROT:Q9YEL2; DDBJ:AP0000060; NID:g5104188; PIDN:BAW79534.1; PID:g51
A:Experimental source: strain K1
C:Genetics:
A:Gene: APE0566
C:Superfamily: proline-rich protein

Query Match 100.0%; Score 35; DB 2; Length 215;
Best Local Similarity 100.0%; Pred. No. 55;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 57 PPHG 61

RESULT 12
B87440
Transcription regulator, TetR family [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: B87440
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.I
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: B87440
A>Status: preliminary
A:Molecule type: DNA

```


Best Local Similarity 100.0%; Pred. No. 81;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 270 PPHG 274

RESULT 18
I48338

C:Species: Mus musculus (house mouse)
C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
C:Accession: I48338
R:Buck, V.; Allen, K.E.; Sorensen, T.; Bybee, A.; Hijmans, E.M.; Voorhoeve, P.M.; Bernat
Oncogene 11, 31-38, 1995
A:Title: Molecular and functional characterisation of E2F-5, a new member of the E2F fam
A:Reference number: I48338; MUID:95349934; PMID:7542760
A:Accession: I48338
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-335 <RES>
A:Cross-references: UNIPROT:Q61502; EMBL:X86925; NID:g806571; PIDN:CAA60508.1; PID:g8065
C:Genetics:
A:Gene: e2f-5

Query Match 100.0%; Score 35; DB 2; Length 335;
Best Local Similarity 100.0%; Pred. No. 86;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 20 PPHG 24

RESULT 19
T26449

hypothetical protein Y113G7B.23 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T26449
R:Lennard, N.
submitted to the EMBL Data Library, September 1999
A:Reference number: Z20215
A:Accession: T26449
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-366 <WIL>
A:Cross-references: UNIPROT:Q8U2M6; EMBL:AL110477; NID:e1542121; PIDN:CAB54337.1; CESP:Y
C:Genetics:
A:Experimental source: clone Y113G7B
A:Gene: CESP:Y113G7B.23
A:Introns: 318/1

Query Match 100.0%; Score 35; DB 2; Length 366;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 330 PPHG 334

RESULT 20
G84769

hypothetical protein At2g35530 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84769
R:Lin, X.; Kaul, S.; Roundley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J
Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84769
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-368 <STO>
A:Cross-references: UNIPROT:O82288; GB:AE002093; NID:g3608135; PIDN:AAC36168.1; GSPDB:GN
C:Genetics:
A:Gene: At2g35530
A:Map position: 2
C:Superfamily: fava bean G-box-binding protein; fos/jun DNA-binding domain homology

Query Match 100.0%; Score 35; DB 2; Length 368;
Best Local Similarity 100.0%; Pred. No. 94;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 54 PPHG 58

RESULT 21
G87723

protein R06A10.4 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C:Accession: G87723
R:Anonymous, The C. elegans Sequencing Consortium.
Science 282, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biology
A:Reference number: A75000; MUID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_eleg
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: G87723
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-391 <STO>
A:Cross-references: UNIPROT:O44747; GB:chr_I; PIDN:AAB96730.1; PID:g2773215; GSPDB:GN0000
C:Genetics:
A:Gene: R06A10.4
A:Map position: 1

Query Match 100.0%; Score 35; DB 2; Length 391;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 40 PPHG 44

RESULT 22
T07967

mandelonitrile lyase (EC 4.1.2.10) - flax
N:Alternate names: hydroxynitrile lyase
C:Species: Linum usitatissimum (flax)
C:Date: 21-May-1999 #sequence_revision 21-May-1999 #text_change 09-Jul-2004
C:Accession: T07967
R:Breithaupt, H.M.; Boenigk, W.; Pohl, M.; Kula, M.R.
submitted to the EMBL Data Library, September 1997
A:Description: Cloning and overexpression of (R)-hydroxynitrile lyase from Linum usitatiss
A:Reference number: Z16253
A:Accession: T07967
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-422 <BRE>
A:Cross-references: UNIPROT:O22574; EMBL:AF024588; NID:g25655287; PIDN:AAB81956.1; PID:g2
C:Genetics:
A:Gene: HNL
C:Function:
A:Description: catalyzes reversible addition of HCN to aliphatic ketones and aldehydes
C:Superfamily: alcohol dehydrogenase; long-chain alcohol dehydrogenase homology
C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match 100.0%; Score 35; DB 2; Length 422;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
DB 190 PPPHG 194

RESULT 23
T32632
hypothetical protein F15E6.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T32632
R:Miller, N.; Stellyes, L.; Bradshaw, H.; Keppler, D.
submitted to the EMBL Data Library, December 1997
A:Description: The sequence of C. elegans cosmid F15E6.
A:Reference number: Z21202
A:Accession: T32632
A:Status: preliminary; translated from GB/EMBL/DBSJ
A:Molecule type: DNA
A:Residues: 1-429 <ML>
A:Cross-references: UNIPROT:O44497; EMBL:AF038614; PIDN:AAB92062.1; GSPDB:GN00022; CBSP:
A:Experimental source: strain Bristol N2; clone F15E6
C:Genetics:
A:Gene: CBSP:F15E6.2
A:Map position: 4
A:Introns: 196/3; 241/3; 338/1

Query Match 100.0%; Score 35; DB 2; Length 429;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
DB 369 PPPHG 373

RESULT 24
S19010
homeotic protein PBX3a - human
N:Alternate names: homeobox protein; pre-B-cell leukemia transcription factor 3
C:Species: Homo sapiens (man)
C>Date: 29-Nov-1993 #sequence_revision 10-Nov-1995 #text_change 16-Aug-2004
C:Accession: S19010
R:Montca, K.; Galili, N.; Nourse, J.; Saltman, D.; Cleary, M.L.
Mol. Cell. Biol. 11, 6149-6157, 1991
A:Title: PBX2 and PBX3, new homeobox genes with extensive homology to the human proto-on
A:Reference number: S19009; MUID:92049345; PMID:1682799
A:Accession: S19010
A:Status: preliminary; nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 1-434 <MON>
A:Cross-references: UNIPROT:P40426; EMBL:X59841; NID:G35314; PIDN:CAA42502.1; PID:G35315
C:Genetics:
A:Gene: GDB:PBX3
A:Cross-references: GDB:125353; OMIM:176312
A:Map position: 9q33-q9q34
C:Superfamily: homeobox homology
C:Keywords: DNA binding; homeobox; nucleus; transcription regulation
F:236-295/Domain: homeobox homology <HOX>

Query Match 100.0%; Score 35; DB 2; Length 434;
Best Local Similarity 100.0%; Pred. No. 1.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
DB 29 PPPHG 33

RESULT 25
F81046
hypothetical protein NMB1759 - Neisseria meningitidis (strain MC58 serogroup B)
C:Species: Neisseria meningitidis
C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C:Accession: F81046
R:Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.A.;
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.;
xi, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scariato, V.; Masignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A:Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; Ver
A:Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A:Reference number: AB1000; MUID:20175755; PMID:10710307
A:Accession: F81046
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-454 <TET>
A:Cross-references: UNIPROT:Q9JY39; GB:AE002525; GB:AE002098; NID:G7227004; PIDN:AAF42101
A:Experimental source: serogroup B, strain MC58
C:Genetics:
A:Gene: NMB1759
C:Superfamily: Neisseria meningitidis hypothetical protein NMB1759

Query Match 100.0%; Score 35; DB 2; Length 454;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
DB 234 PPPHG 238

RESULT 26
B69458
sulfate adenylyltransferase (sat) homolog - Archaeoglobus fulgidus
C:Species: Archaeoglobus fulgidus
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 21-Jul-2000
R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson,
.; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.;
Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.
Nature 390, 364-370, 1997
A:Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.B.
Smith, H.O.; Woese, C.R.; Venter, J.C.
A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archaeo
A:Reference number: A69250; MUID:98049343; PMID:9389475
A:Accession: B69458
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-459 <KLE>
A:Cross-references: GB:AE000988; GB:AE000782; NID:G2689311; PIDN:AAB89581.1; PID:G264888
C:Superfamily: sulfate adenylyltransferase met3-1; sulfate adenylyltransferase homology

Query Match 100.0%; Score 35; DB 2; Length 459;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
DB 10 PPPHG 14

RESULT 27
G02738
F8EAC-4 - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 16-Aug-2004
C:Accession: G02738
R:Enserback, S.
submitted to the EMBL Data Library, June 1996
A:Reference number: H01646
A:Accession: G02738
A:Status: preliminary; translated from GB/EMBL/DBSJ

A:Molecule type: mRNA
A:Residues: 1-465 <ENE>
A:Cross-references: UNIPROT:Q16676; EMBL:U59832; NID:G1399238; PID:G1399239
C:Genetics:
A:Gene: FREAC-4
C:Superfamily: fork head DNA-binding domain homology
F:125-216/Domain: fork head DNA-binding domain homology <FHD>

Query Match 100.0%; Score 35; DB 2; Length 465;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 313 PPPHG 317

RESULT 28

A44461
voltage-dependent calcium channel beta-1 chain, M isoform - human
C:Species: Homo sapiens (man)
C:Date: 10-Jun-1993 #sequence_revision 18-Nov-1994 #text_change 01-Dec-2000
C:Accession: A44461; I38001; S31919
R:Powers, P.A.; Liu, S.; Hogan, K.; Gregg, R.G.
J. Biol. Chem. 267, 22967-22972, 1992
A:Title: Skeletal muscle and brain isoforms of a beta-subunit of human voltage-dependent
A:Reference number: A44461; MUID:93054616; PMID:1385409
A:Accession: A44461
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-475 <POW>
A:Cross-references: GB:M92301
A:Experimental source: skeletal muscle
A:Note: sequence extracted from NCBI backbone (NCBIP:118128); sequence inconsistent with
R:iles, D.E.; Segers, B.; Sengers, R.C.; Monsieus, K.; Heytens, L.; Halseall, P.J.; Hopk
Hum. Mol. Genet. 2, 863-868, 1993
A:Title: Genetic mapping of the beta 1- and gamma-subunits of the human skeletal muscle
thermia susceptibility.
A:Reference number: I38001; MUID:93372845; PMID:8395940
A:Accession: I38001
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 146-183 <ILE>
A:Cross-references: EMBL:Z21725; NID:938562; PIDN:CAA79824.1; PID:938563
A:Note: submitted to the EMBL Data Library, February 1993
C:Superfamily: human voltage-dependent calcium channel beta chain

Query Match 100.0%; Score 35; DB 2; Length 475;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 224 PPPHG 228

RESULT 29

T20482
hypothetical protein F01G10.10 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C:Accession: T20482
R:Hembry, C.
submitted to the EMBL Data Library, October 1996
A:Reference number: Z19280
A:Accession: T20482
A>Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-486 <WIL>
A:Cross-references: UNIPROT:O17768; EMBL:Z81055; PIDN:CAB02898.1; GSPDB:GN00022; CBSP:FO
A:Experimental source: clone F01G10
C:Genetics:
A:Gene: CBSP:F01G10.10

A:Map position: 4
A:Introns: 36/3; 73/2; 114/2; 154/1; 190/3; 211/3; 287/1; 364/3; 398/3; 449/2
Query Match 100.0%; Score 35; DB 2; Length 486;
Best Local Similarity 100.0%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 283 PPPHG 287

RESULT 30

JC6203
SP8 binding protein homolog - cucumber
C:Species: Cucumis sativus (cucumber)
C:Date: 11-Apr-1997 #sequence_revision 09-May-1997 #text_change 09-Jul-2004
C:Accession: JC6203
R:Kim, D.J.; Smith, S.M.; Leaver, C.J.
Gene 185, 265-269, 1997
A:Title: A cDNA encoding a putative SP1-type DNA-binding protein from cucumber.
A:Reference number: JC6203; MUID:97208883; PMID:9055825
A:Accession: JC6203
A:Molecule type: mRNA
A:Residues: 1-509 <KIM>
A:Cross-references: UNIPROT:Q39658; GB:I44134; NID:927024; PIDN:AC37515.1; PID:927025
C:Comment: This protein is a DNA-binding protein consisting of a monomer. It is involved
C:Genetics:
A:Gene: sz71
C:Keywords: DNA binding

Query Match 100.0%; Score 35; DB 2; Length 509;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 226 PPPHG 230

RESULT 31

T49422
RAD57 related protein [imported] - Neurospora crassa
N:Alternate names: protein B17C10.30
C:Species: Neurospora crassa
C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 09-Jul-2004
C:Accession: T49422
R:Schulte, U.; Aign, V.; Hoheisel, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatara,
submitted to the Protein Sequence Database, May 2000
A:Reference number: Z25022
A:Accession: T49422
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-516 <SCH>
A:Cross-references: UNIPROT:Q9P6E6; EMBL:AL355926; GSPDB:GN00116; NCSP:B17C10.30
A:Experimental source: BAC clone B17C10; strain OR74A
C:Genetics:
A:Gene: NCSP:B17C10.30
A:Map position: 6
A:Introns: 31/3

Query Match 100.0%; Score 35; DB 2; Length 516;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
|||
Db 158 PPPHG 162

RESULT 32

S61920
B-alpha pheromone-receptor 2 - bracket fungus (Schizophyllum commune) (fragment)

C;Species: Schizophyllum commune
C;Date: 23-Jul-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S61920
R;Wendland, J.; Vailancourt, L.J.; Hegner, J.; Lengeler, K.B.; Laddison, K.J.; Specht, EMO J. 14, 5271-5278, 1995
A;Title: The mating-type locus Balphal of Schizophyllum commune contains a pheromone receptor gene
A;Reference number: S61919; MUID:96080162; PMID:7489716
A;Accession: S61920
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-518 <WEN>
A;Cross-references: UNIPROT:Q05659; EMBL:X91168; NID:gl143553; PIDN:CAA62595.1; PID:gl143553
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1995
C;Genetics:
A;Introns: 71/1; 108/3; 131/1; 177/2; 320/3

Query Match 100.0%; Score 35; DB 2; Length 518;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 203 PPPHG 207

RESULT 33
I65767
L-type voltage-gated calcium channel B subunit - human
C;Species: Homo sapiens (man)
C;Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 18-Aug-2000
C;Accession: I65767
R;Collin, T.; Wang, J.; Nargeot, J.; Schwartz, A.
Circ. Res. 72, 1337-1344, 1993
A;Title: Molecular cloning of three isoforms of the L-type voltage-dependent calcium channel alpha1C
A;Reference number: I52859; MUID:93265672; PMID:7916667
A;Accession: I65767
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-522 <RES>
A;Cross-references: GB:I06112; NID:gl87018; PIDN:AAA36169.1; PID:gl87019
C;Superfamily: human voltage-dependent calcium channel beta chain

Query Match 100.0%; Score 35; DB 2; Length 522;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 248 PPPHG 252

RESULT 34
A41347
calcium channel protein beta chain, dihydropyridine-sensitive, skeletal muscle - rabbit
C;Species: Oryctolagus cuniculus (domestic rabbit)
C;Date: 03-Apr-1992 #sequence_revision 03-Apr-1992 #text_change 09-Jul-2004
C;Accession: A41347
R;Ruth, P.; Roehrkasten, A.; Biel, M.; Bosse, E.; Regulla, S.; Meyer, H.E.; Flockerzi, V.
Science 245, 1115-1118, 1999
A;Title: Primary structure of the beta subunit of the DHP-sensitive calcium channel from rabbit skeletal muscle
A;Reference number: A41347; MUID:89368946; PMID:2549640
A;Accession: A41347
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-524 <RUT>
A;Cross-references: UNIPROT:P19517; GB:M25817
C;Superfamily: human voltage-dependent calcium channel beta chain

Query Match 100.0%; Score 35; DB 2; Length 524;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 249 PPPHG 253

RESULT 35
T00677
hypothetical protein At2g43970 [imported] - Arabidopsis thaliana
N;Alternate names: hypothetical protein F6E13.10
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 01-Feb-1999 #sequence_revision 01-Feb-1999 #text_change 09-Jul-2004
C;Accession: T00677; G84872
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.; Rounsley, S.D.; Kaul, S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Description: Arabidopsis thaliana chromosome II BAC F6E13 genomic sequence.
A;Reference number: Z14180
A;Accession: T00677
A;Status: translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-529 <ROU>
A;Cross-references: UNIPROT:O80567; EMBL:AC004005; NID:g3212846; PID:g3212854
A;Experimental source: cultivar Columbia
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84872
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-529 <STO>
A;Cross-references: GB:AE002093; NID:g3212854; PIDN:AAC23405.1; GSPDB:GNO0139
C;Genetics:
A;Gene: F6E13.10; At2g43970
A;Map position: 2
A;Introns: 200/3; 228/1; 259/3; 283/3; 310/2; 344/3; 367/3; 421/3

Query Match 100.0%; Score 35; DB 2; Length 529;
Best Local Similarity 100.0%; Pred. No. 1.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5
|||||
Db 76 PPPHG 80

RESULT 36
G84825
probable CCH-type zinc finger protein [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C;Accession: G84825
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A;Reference number: A84420; MUID:20083487; PMID:10617197
A;Accession: G84825
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-597 <STO>
A;Cross-references: UNIPROT:Q9XEE6; GB:AE002093; NID:g6598933; PIDN:AAF18728.1; GSPDB:GNC
C;Genetics:
A;Gene: At2g40140
A;Map position: 2

Query Match 100.0%; Score 35; DB 2; Length 597;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPPHG 5

```
Db          456 PPPHG 460
|||||
RESULT 37
E86194
hypothetical protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E86194
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.
C.R.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maity, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E86194
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-614 <STO>
A:Cross-references: UNIPROT:Q9LNF0; GB:AE005172; NID:98810458; PIDN:AAF80119.1; GSPDB:GN
C:Genetics:
A:Map position: 1
Query Match          100.0%; Score 35; DB 2; Length 614;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db          46 PPPHG 50
|||||
RESULT 38
S48424
hypothetical protein YIL055c - yeast (Saccharomyces cerevisiae)
C:Species: Saccharomyces cerevisiae
C>Date: 02-Dec-1994 #sequence_revision 02-Dec-1994 #text_change 09-Jul-2004
C:Accession: S48424
R:Smith, V.
submitted to the EMBL Data Library, September 1994
A:Reference number: S48407
A:Accession: S48424
A:Molecule type: DNA
A:Residues: 1-627 <SMI>
A:Cross-references: UNIPROT:P40523; GB:Z47047; EMBL:Z38060; NID:g603997; PID:g763291; GS
C:Genetics:
A:Gene: MIPS:YIL055c
A:Cross-references: SGD:S0001317
A:Map position: 9L
Query Match          100.0%; Score 35; DB 2; Length 627;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db          222 PPPHG 226
|||||
RESULT 39
C87048
probable ABC transporter, ATP-binding component MLI13 [imported] - Mycobacterium leprae
C:Species: Mycobacterium leprae
C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: C87048
R:Cole, S.T.; Eiglmeier, K.; Parkhill, J.; James, K.D.; Thomson, N.R.; Wheeler, P.R.; Ho
R.; Davies, R.M.; Devlin, K.; Duthoy, S.; Feltwell, T.; Fraser, A.; Hamlin, N.; Holroyd,
```

```
eam, M.A.; Rutherford, K.M.
Nature 409, 1007-1011, 2001
A:Authors: Rutter, S.; Seeger, K.; Simon, S.; Simmonds, M.; Skelton, J.; Squares, R.; Sq
A:Title: Massive gene decay in the leprosy bacillus.
A:Reference number: A86909; MUID:21128732; PMID:11234002
A:Accession: C87048
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-629 <STO>
A:Cross-references: UNIPROT:Q9CC89; GB:AL450380; NID:g13093097; PIDN:CAC31494.1; GSPDB:G
C:Genetics:
A:Gene: MLI13
C:Superfamily: Escherichia coli ABC transporter mdla; ATP-binding cassette homology
Query Match          100.0%; Score 35; DB 2; Length 629;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db          382 PPPHG 386
|||||
RESULT 40
A49585
Na+ channel protein, amiloride-sensitive - human
C:Species: Homo sapiens (man)
C>Date: 16-Feb-1996 #sequence_revision 16-Feb-1996 #text_change 09-Jul-2004
C:Accession: A49585; I51911
R:Voilley, N.; Lingueglia, E.; Champigny, G.; Mattei, M.G.; Waldmann, R.; Lazdunski, M.;
Proc. Natl. Acad. Sci. U.S.A. 91, 247-251, 1994
A:Title: The lung amiloride-sensitive Na+ channel: biophysical properties, pharmacology,
A:Reference number: I51911; MUID:94105144; PMID:8278374
A:Accession: A49585
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-669 <RES>
A:Cross-references: UNIPROT:P37088; EMBL:X76180; NID:g452649; PIDN:CAA53773.1; PID:g4526;
R:McDonald, F.J.; Snyder, P.M.; McCray, P.B.
Am. J. Physiol. 266, L728-L734, 1994
A:Title: Cloning, expression, and tissue distribution of a human amiloride-sensitive Na+
A:Reference number: I51911; MUID:94295729; PMID:8023962
A:Accession: I51911
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-669 <RE2>
A:Cross-references: GB:L29007; NID:g493125; PIDN:AAA21813.1; PID:g493605
C:Genetics:
A:Map position: 12p13
C:Superfamily: human amiloride-sensitive sodium channel protein; fibronectin type I repe
Query Match          100.0%; Score 35; DB 2; Length 669;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPPHG 5
Db          195 PPPHG 199
|||||
Search completed: June 15, 2005, 14:22:51
Job time : 19.5 secs
```

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 15, 2005, 13:53:11 ; Search time 83.5 Seconds
(without alignments)
30.663 Million cell updates/sec

Title: US-10-074-225A-10
Perfect score: 35
Sequence: 1 PPHG 5

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : UniProt Q3:*
1: uniprot_gprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Query	ID	Description
1	35	100.0	33	002832	Gorilla gor
2	35	100.0	53	Q9V8Y1	Drosophila
3	35	100.0	62	Q6UB47	hyaloperono
4	35	100.0	72	PN2D	LITSE
5	35	100.0	80	Q9VPZ1	Drosophila
6	35	100.0	93	Q94J14	oryza sativ
7	35	100.0	94	Q62CG2	burkholderi
8	35	100.0	94	Q63KL8	burkholderi
9	35	100.0	108	Q655P3	oryza sativ
10	35	100.0	111	Q9YCY2	aeropyrum p
11	35	100.0	119	Q67867	hepatitis b
12	35	100.0	134	Q7YWR5	caenorhabdi
13	35	100.0	135	Q7VVE3	bordelella
14	35	100.0	135	Q7WBB4	bordelella
15	35	100.0	135	Q7WMT3	bordelella
16	35	100.0	147	SMR1_MOUSE	mus musculu
17	35	100.0	151	Q8N1P8	homo sapien
18	35	100.0	152	Q8ABU0	bacteroides
19	35	100.0	164	Q8GRK5	oryza sativ
20	35	100.0	164	Q8GSL3	oryza sativ
21	35	100.0	164	Q657L2	oryza sativ
22	35	100.0	164	Q62HZ3	burkholderi
23	35	100.0	164	Q63RT2	burkholderi
24	35	100.0	165	Q62W55	homo sapien
25	35	100.0	168	Q06065	mycobacteri
26	35	100.0	169	Q07429	mycobacteri
27	35	100.0	169	Q7U2M4	mycobacteri
28	35	100.0	170	Q07610	rattus norv
29	35	100.0	173	Q62063	caenorhabdi
30	35	100.0	173	Q62789	sus scrofa
31	35	100.0	173	Q8LLM4	tritimum ae

RESULT 1
002832 ID 002832 PRELIMINARY; PRT; 33 AA.
AC 002832;
DT 01-JUL-1997 (Tremblrel. 04, Created)
DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE Huntingtin protein (fragment).
GN Name=IT15;
OS Gorilla gorilla (gorilla).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Gorilla.
OX NCBI_TaxID=9593;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96326790; PubMed=8766138;
RA Pecheux C., Gall A.L., Kaplan J.C., Dode C.;
RT "Sequence analysis of the CAG triplet repeats region in the Huntington
RT disease gene (IT15) in several mammalian species."
RL Ann. Genet. 39:81-86(1996).
DR EMBL; S83377; AAB50771.1; -.
FT NON TER 33
SQ SEQUENCE 33 AA; 3393 MW; 44F558CE5636722A CRC64;

Query Match 100.0%; Score 35; DB 2; Length 33;
Best Local Similarity 100.0%; Pred. No. 46;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 PPHG 5
Db 5 PPHG 9

RESULT 2
002832 ID 002832 PRELIMINARY; PRT; 53 AA.
AC 002832;
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE CG15126-PA.
GN ORFNames=CG15126;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

Q9LF59 arbidopsis
Q52724 rhodococcus
Q8LEP5 arbidopsis
Q9C4Z8 arbidopsis
Q28261 canis famil
Q8ZY20 pyrobaculum
Q67U43 oryza sativ
Q8ND66 homo sapien
Q67IV2 oryza sativ
P04474 rattus norv
Q9VEL2 aeropyrum p
Q9FYN6 oryza sativ
Q8RYZ6 oryza sativ
Q9A825 caulobacter

ALIGNMENTS

CC lesser extent in small granules of hemocytes.
 CC -I- SIMILARITY: Belongs to the peneidin family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.ebi.ac.uk/ebis/announcements/>
 CC or send an email to license@ebi.ac.uk).
 CC -----
 CC EMBL; AY039205; AAK83453.1; -
 DR HSP; P81058; 1UEO.
 DR InterPro: IPR009226; Peneidin.
 DR Pfam: PF05927; Peneidin; 1.
 KW Amidation; Antibiotic; Chitin-binding; Fungicide;
 KW Pyroglutamate carboxylic acid; Signal.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 71 Peneidin-2d.
 FT DOMAIN 29 43 Pro-rich.
 FT DISULFID 45 59 By similarity.
 FT DISULFID 48 66 By similarity.
 FT DISULFID 60 67 By similarity.
 FT MOD_RES 22 22 Pyroglutamate carboxylic acid (By
 FT similarity).
 FT MOD_RES 71 71 Lysine amide (G-72 provides amide group)
 FT (By similarity).
 FT SIGNAL 71 71 7DCE24C8F3D3F182 CRC64;
 SQ SEQUENCE 72 AA; 7717 MW; 70CE24C8F3D3F182 CRC64;
 Query Match 100.0%; Score 35; DB 1; Length 72;
 Best Local Similarity 100.0%; Pred. No. 1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 33 PPPHG 37
 RESULT 5
 ID Q9VPZ1 PRELIMINARY; PRT; 80 AA.
 AC Q9VPZ1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE CG5011-PA.
 GN ORFNames=CG5011;
 OS Drosophila melanogaster (fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.H., Blazell R.G., Champ M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kianos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svitek R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodger, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhao M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster";
 RL Science 287:2185-2195(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svitek R., Tabor P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstein G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RT "Finishing a whole-genome shotgun: Release 3 of the Drosophila
 RT melanogaster euchromatic genome sequence";
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svitek R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RT "The transposable elements of the Drosophila melanogaster euchromatin:
 RT a genomics perspective";
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Milburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Bettencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RT "Annotation of the Drosophila melanogaster euchromatic genome: a
 RT systematic review";
 RL Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003587; AAF51393.1; -
 DR IntAct; Q9VPZ1; -
 DR FlyBase; FBgn0040723; CG5011.
 SQ SEQUENCE 80 AA; 8699 MW; D57E2D6B3443B4BF CRC64;
 Query Match 100.0%; Score 35; DB 2; Length 80;
 Best Local Similarity 100.0%; Pred. No. 1.1e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 PPPHG 5
 Db 33 PPPHG 37

```
Db 12 PPHG 16

RESULT 6
Q94J14
ID Q94J14 PRELIMINARY; PRT; 93 AA.
AC Q94J14;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE P0481E12.34 protein.
GN Name=P0481E12.34;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Wu J., Niumura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikono M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003076; BAB56049.1; -.
DR Gramene; Q94J14; -.
SQ SEQUENCE 93 AA; 9989 MW; 6EDA478C67F82098 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 93;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 85 PPHG 89

RESULT 7
Q62CG2
ID Q62CG2 PRELIMINARY; PRT; 94 AA.
AC Q62CG2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE DNA-binding protein.
GN ORFNames=BMAA0931.
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344.
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Rinning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammed Y., Nelson W.C., Radune D., Romero C.M., Sarria S.,
RA Selengut J., Shamlin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL; CP000011; AAU46112.1; -.
KW DNA-binding.
SQ SEQUENCE 94 AA; 9817 MW; E0DA5022E10EE8FF CRC64;

Query Match 100.0%; Score 35; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 75 PPHG 79

RESULT 8
Q63KL8
ID Q63KL8 PRELIMINARY; PRT; 94 AA.
AC Q63KL8;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative DNA-binding protein.
GN ORFNames=BPSS1343;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Church C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Chervach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshazer D.,
RA Feltwell T., Fraser A., Hance Z., Hauger H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571966; CAH38814.1; -.
KW DNA-binding.
SQ SEQUENCE 94 AA; 9817 MW; E0DA5022E10EE8FF CRC64;

Query Match 100.0%; Score 35; DB 2; Length 94;
Best Local Similarity 100.0%; Pred. No. 1.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 75 PPHG 79

RESULT 9
Q655P3
ID Q655P3 PRELIMINARY; PRT; 108 AA.
AC Q655P3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein B1011A07.27.
GN Name=B1011A07.27;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
```

```

RX PubMed=12447438;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nimura Y., Cheng Z., Negamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh T., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Nachita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki K., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL; AP003722; BAD45462.1; -.
KW Hypothetical protein.
SQ SEQUENCE 108 AA; 11886 MW; 6B93CD83685C0A7 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 108;
Best Local Similarity 100.0%; Pred. No. 1.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 47 PPHG 51

RESULT 10
Q9YCY2 PRELIMINARY; PRT; 111 AA.
AC Q9YCY2;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein APE1130.
GN OrderedLocustNames=APE1130;
OS Aeropyrum pernix.
OS Archaea; Crenarchaeota; Thermoprotei; Desulfurococcales;
OC Desulfurococccaceae; Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]_TaxID=56636;
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX Kwarabayasi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anka A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Mishiina K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudo H. Y.,
RA Yamazaki J., Kushida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT "Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1.";
RL DNA Res. 6:83-101(1999).
DR EMBL; AP000060; BAA80115.1; -.
DR PIR; C72714; C72714.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 111 AA; 12537 MW; 3CE2AE76F92522BB CRC64;

Query Match 100.0%; Score 35; DB 2; Length 111;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 82 PPHG 86

RESULT 11
Q67867 PRELIMINARY; PRT; 119 AA.
ID Q67867

```

```

AC Q67867;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Pre-S1 region (Fragment).
OS Hepatitis B virus.
OC Viruses; Retroviruses; Hepadnaviridae; Orthohepadnavirus.
OX NCBI_TaxID=10407;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=adr;
RX MEDLINE=89183619; PubMed=2928116;
RA Rho H.M., Kim K., Hyun S.W., Kim Y.S.;
RT "The nucleotide sequence and reading frames of a mutant hepatitis B
RT virus adr.";
RL Nucleic Acids Res. 17:2124-2124(1989).
DR EMBL; X14193; CAA32406.1; -.
GO; GO:0016032; P:Viral life cycle; IEA.
DR InterPro; IPR000349; Hepvir_surflg.
DR Pfam; PF00695; VMSA; I.
KW Antigen.
FT NON TER 119 119
SQ SEQUENCE 119 AA; 12553 MW; 8BE9E93CD66F1E6D CRC64;

Query Match 100.0%; Score 35; DB 2; Length 119;
Best Local Similarity 100.0%; Pred. No. 1.7e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 68 PPHG 72

RESULT 12
Q7YWR5 PRELIMINARY; PRT; 134 AA.
ID Q7YWR5
AC Q7YWR5;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein Y40B1A.5.
GN ORFNames=Y40B1A.5;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]_TaxID=6239;
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C.elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Harris B.R.;
RL Submitted (DEC-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL034392; CAE17989.1; -.
DR WormBase; WBGene00012736; Y40B1A.5.
DR WormPep; Y40B1A.5; CE35082.
KW Hypothetical protein.
SQ SEQUENCE 134 AA; 14532 MW; E1CFF526E97C297D CRC64;

Query Match 100.0%; Score 35; DB 2; Length 134;
Best Local Similarity 100.0%; Pred. No. 1.9e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 88 PPHG 92

```


28-FEB-2003 (Rel. 41, Created)
28-FEB-2003 (Rel. 41, Last sequence update)
05-JUL-2004 (Rel. 44, Last annotation update)
Submaxillary gland androgen regulated protein 1 precursor (Salivary protein MSG1).
GN Name=Smrl; Synonyms=MSG1;
OS Mus musculus (Mouse).
OC Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Salivary gland;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
RA Klautner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettaman M., Madan A.C., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Rutherford J.S.N., Krzywicki M.I., Skalska U., Smalhus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May play a role in protection or detoxification.
CC -!- SUBCELLULAR LOCATION: Secreted.
CC -!- TISSUE SPECIFICITY: Secreted into saliva by submaxillary gland.
CC -!- SIMILARITY: Belongs to the PROL1/PROL3 family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X71629; CAA50636.1; -;
DR EMBL; BC031806; AAH31806.1; -;
DR EMBL; BC031921; AAH31921.1; -;
DR EMBL; BC034553; AAH34553.1; -;
DR EMBL; BC055857; AAH55857.1; -;
DR FGI; 148669; S37485.
DR MGD; MGI:102763; Smrl.
KW Repeat; Signal.
FT SIGNAL 1 22 Potential.
FT CHAIN 23 147 Submaxillary gland androgen regulated
FT FT protein 1.
FT FT 3 X 12 AA tandem repeats of G-P-G-I-G-R-
FT FT P-[HP]-P-P-P-[PF].
FT FT 1.
FT FT 43 54 1.
FT FT REPEAT 55 66 2.
FT FT REPEAT 67 78 3.
FT FT 73 84 Poly-Pro.
FT FT DOMAIN

SQ SEQUENCE 147 AA; 15544 MW; 07D64E22F1BF58C2 CRC64;
Query Match 100.0%; Score 35; DB 1; Length 147;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPHG 5
Db 39 PPHG 43
RESULT 17
Q8N1P8 PRELIMINARY; PRT; 151 AA.
AC Q8N1P8;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2002 (TRENBLrel. 22, Last annotation update)
DE Hypothetical protein FLJ38012.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=rongue;
RX PubMed=14702039; DOI=10.1038/ngl1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakatsuki A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Moniyan H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human cDNAs.";
RL Nat. Genet. 36:40-45(2004).
DR EMBL; AK095331; BAC04532.1; -;
SQ SEQUENCE 151 AA; 16189 MW; 091EADD64FECE89 CRC64;
Query Match 100.0%; Score 35; DB 2; Length 151;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 PPHG 5
Db 19 PPHG 23
RESULT 18
Q8ABU0 PRELIMINARY; PRT; 152 AA.
ID Q8ABU0

```

AC Q8ABU0;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN OrderedLocustNames=BT0020;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]_TaxID=818;
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RA MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RX Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AB016926; AA075127.1; -.
KW Complete proteome.
SQ SEQUENCE 152 AA; 17665 MW; 1991F9218509F8D9 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 152;
Best Local Similarity 100.0%; Pred. No. 2.1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 143 PPPHG 147

RESULT 19
Q8GRK5 PRELIMINARY; PRT; 164 AA.
AC Q8GRK5;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ112_E08.108 (Hypothetical protein
DE OJ112_E08.111) (Hypothetical protein OJ112_E08.113) (Hypothetical
DE protein OJ1720_F04.121) (Hypothetical protein OJ1720_F04.124).
GN Name=OJ112_E08.108;
GN Synonyms=OJ112_E08.111, OJ112_E08.113, OJ1720_F04.121,
GN OJ1720_F04.124;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ112_E08.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 49 PPPHG 53

RESULT 20
Q8GSL3 PRELIMINARY; PRT; 164 AA.
AC Q8GSL3;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein OJ112_E08.110 (Hypothetical protein
DE OJ1720_F04.123).
GN Name=OJ112_E08.110; Synonyms=OJ1720_F04.123;
GN Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]_TaxID=39947;
RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 7, BAC
RT clone:OJ112_E08.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [2]

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 49 PPPHG 53

RESULT 21
Q857L2 PRELIMINARY; PRT; 164 AA.
AC Q857L2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein P0468H06.16;
GN Name=P0468H06.16;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Niimura Y., Cheng Z., Nagamura Y., Antonio B.A., Kanamori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,

Query Match 100.0%; Score 35; DB 2; Length 164;

```

```

RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
RA Ikano M., Itoh S., Itoh T., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machata K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagaaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yanagata H., Yanane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.;
RT "The genome sequence and structure of rice chromosome 1.";
RL Nature 420:312-316(2002).
DR EMBL: AP003075; BAD45005.1; -.
KW Hypothetical protein.
SQ SEQUENCE 164 AA; 18134 MW; 6CF81BD88C3C0C7 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 135 PPPHG 139

RESULT 22
Q62HZ3
ID Q62HZ3 PRELIMINARY; PRT; 164 AA.
AC Q62HZ3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE TonB domain protein.
GN ORFNames=BMA2105;
OS Burkholderia mallei ATCC 23344.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=243160;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 23344.
RA Nierman W.C., DeShazer D., Kim H.S., Tettelin H., Nelson K.E.,
RA Feldblyum T., Ulrich R.L., Ronning C.M., Brinkac L.M., Daugherty S.C.,
RA Davidson T.D., Deboy R.T., Dimitrov G., Dodson R.J., Durkin A.S.,
RA Gwinn M.L., Haft D.H., Khouri H., Kolonay J.F., Madupu R.,
RA Mohammad Y., Nelson W.C., Radune D., Romero C.M., Sarrisa S.,
RA Selengut J., Shambin C., Sullivan S.A., White O., Yu Y., Zafar N.,
RA Zhou L., Fraser C.M.;
RT "Structural flexibility in the Burkholderia mallei genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14247-14251(2004).
DR EMBL: CP000010; AAU50017.1; -.
SQ SEQUENCE 164 AA; 17543 MW; 3B29EF2562073ABB CRC64;

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 34 PPPHG 38

RESULT 23
Q63RT2
ID Q63RT2 PRELIMINARY; PRT; 164 AA.
AC Q63RT2;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative lipoprotein.
GN ORFNames=BPSL2589;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;

```

```

OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K., I.R.,
RA Bentley S.D., Sebaihia M., Thomson N.R., Bason N., Beacham I.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crossset B., Davis P., DeShazer D.,
RA Fellwell T., Fraser A., Hance Z., Hauser H., Holroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitz E., Rutherford K., Tumapa S., Vesaratchavest M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchavest M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL: BX571965; CH36597.1; -.
KW Lipoprotein.
SQ SEQUENCE 164 AA; 17559 MW; 3B360F257DE73ABB CRC64;

Query Match 100.0%; Score 35; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 34 PPPHG 38

RESULT 24
Q6ZW55
ID Q6ZW55 PRELIMINARY; PRT; 165 AA.
AC Q6ZW55;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein FLJ41588.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Tongue;
RA Ota T., Nakagawa S., Senoh A., Mizuguchi H., Inagaki H., Sugiyama T.,
RA Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S., Yamamoto J.,
RA Isono Y., Kawai-Hio Y., Saito K., Nishikawa T., Kimura K.,
RA Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H., Kanda K.,
RA Wagatsuma M., Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A.,
RA Sugiyama A., Kawakami B., Suzuki Y., Sugano S., Nagahari K.,
RA Masuho Y., Nagai K., Isogai T.;
RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AK123582; BAC85651.1; -.
SQ SEQUENCE 165 AA; 17648 MW; 9CADD2C8DFCB90F CRC64;

Query Match 100.0%; Score 35; DB 2; Length 165;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
DB 33 PPPHG 37

RESULT 25
O06065
ID O06065 PRELIMINARY; PRT; 168 AA.
AC O06065;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)

```


Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 132 PPHG 136

RESULT 28

Q07610 PRELIMINARY; PRT; 170 AA.
AC Q07610; 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Proline-rich proteoglycan.
GN Names=PRGL;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Rodentia; Sciurognathi; Muridae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Parotid;
RX MEDLINE=93388626; PubMed=8376404;
RA Castle A.M., Castle J.D.;
RT "Novel secretory proline-rich proteoglycans from rat parotid. Cloning
and characterization by expression in AtT-20 cells.";
RL J. Biol. Chem. 268:20490-20496(1993).
DR EMBL; LI7317; AAA03073.1; -.
DR PIR; A48013; A48013.
SQ SEQUENCE 170 AA; 17840 MW; 85AF34C0B45AA03 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 170;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 80 PPHG 84

RESULT 29

O62063 PRELIMINARY; PRT; 173 AA.
AC O62063; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DE Hypothetical protein C16D6.1.
GN ORFNames=C16D6.1;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=99069613; PubMed=9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
investigating biology.";
RL Science 282:2012-2018(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Gardner A.E.;
RL Submitted (NOV-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; 281472; CAB03889.1; -.
DR PIR; T19341; T19341.
DR WormBase; WBGene0007634; C16D6.1.
DR WormPep; C16D6.1; CE08230.
KW Hypothetical protein.
SQ SEQUENCE 173 AA; 19001 MW; 16D04C1BBDF76265 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 72 PPHG 76

RESULT 30

O62789 PRELIMINARY; PRT; 173 AA.
AC O62789; 01-AUG-1998 (TREMBlrel. 07, Created)
DT 01-AUG-1998 (TREMBlrel. 07, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Collagen VIII (Fragment).
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Euthera; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Aortic smooth muscle;
RA Reichenberg S., Plenz G., Robenek H.;
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF054891; AAC08999.1; -.
DR HSSP; Q00780; 1091.
DR InterPro; IPR001073; Clq.
DR InterPro; IPR008983; TNF_like.
DR Pfam; PF00386; Clq; 1.
DR PRINTS; PR00007; COMPLEMENTC1Q.
DR SMART; SM00110; Clq; 1.
DR PROSITE; PS01113; Clq; 1.
KW Collagen.
FT NON_TER 1
FT NON_TER 173
SQ SEQUENCE 173 AA; 18973 MW; 03489B6FBAAB6CAD CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;

Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPHG 5
Db 7 PPHG 11

RESULT 31

Q8LLM4 PRELIMINARY; PRT; 173 AA.
AC Q8LLM4; 01-OCT-2002 (TREMBlrel. 22, Created)
DT 01-OCT-2002 (TREMBlrel. 22, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Thaumatin-like protein 1.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poaceae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RA Vrch Bi I., Kolb F., Boze L., Domier L.;
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF389883; AAM69454.1; -.
DR HSSP; P33679; 1DU5.
DR InterPro; IPR001938; Thaumatin.
DR Pfam; PF00314; Thaumatin; 1.
DR PRINTS; PR00347; THAUMATIN.
DR ProDom; PD001321; Thaumatin; 1.
DR PROSITE; PS00316; THAUMATIN; 1.

```
SQ SEQUENCE 173 AA; 17646 MW; EC172D12ACD944D3 CRC64;
Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 145 PPPHG 149

RESULT 32
Q9LF59 PRELIMINARY; PRT; 173 AA.
AC Q9LF59;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Glycine/proline-rich protein (At5g17650).
GN Name=K10A8.130;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi6.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Sato S., Nakamura Y., Kaneko T., Kato T., Asamizu E., Kotani H.,
RA Tabata S., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Kim C.J., Chen H., Cheuk R., Shinn P., Carninci P., Hayashizaki Y.,
RA Ishida J., Kamiya A., Kawai J., Narusaka M., Sakurai T., Satou M.,
RA Seki M., Shinozaki K., Ecker J.R.;
RA Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
RL EMBL; ALJ391151; CAC01909.1; -.
DR EMBL; BT012163; AAS76258.1; -.
DR PIR; T51469; T51469.
SQ SEQUENCE 173 AA; 18536 MW; E3510947AA98BC0A CRC64;

Query Match 100.0%; Score 35; DB 2; Length 173;
Best Local Similarity 100.0%; Pred. No. 2.4e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 49 PPPHG 53

RESULT 33
Q52724 PRELIMINARY; PRT; 175 AA.
AC Q52724;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Rhodococcus corallinus (NRRL 15444B) N-ethylmaleline chlorohydrolase
DE (trxA) (Fragment).
OS Rhodococcus corallinus.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Corynebacterineae; Gordoniaceae; Gordonia.
OX NCBI_TaxID=36822;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=960111356; PubMed=7592318;
RA Shao Z.Q., Seffens W., Mulbry W., Behki R.M.;
RT "Cloning and expression of the s-triazine hydrolase gene (trxA) from
RT Rhodococcus corallinus and development of Rhodococcus recombinant
```

```
RT strains capable of dealkylating and dechlorinating the herbicide
RT atrazine.";
RT J. Bacteriol. 177:5748-5755(1995).
DR EMBL; L16534; AAA90930.1; -.
DR PIR; T46665; T46665.
DR HSP; Q12178; IPEO.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0008270; F:zinc ion binding; IEA.
KW Hydrolase.
FT NON TER
SQ SEQUENCE 175 AA; 19406 MW; 881A0E9059A62528 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 175;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 129 PPPHG 133

RESULT 34
Q8LEP5 PRELIMINARY; PRT; 176 AA.
AC Q8LEP5;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative glycine and proline-rich protein.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi6.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22089475; PubMed=12093376;
RA Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,
RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
RT "Full-length messenger RNA sequences greatly improve genome
RT annotation.";
RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
RN [2]
RP SEQUENCE FROM N.A.
RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
RA Feldmann K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY085309; AAM62540.1; -.
DR InterPro; IPR006031; XYPX.
DR Pfam; PF02162; XYPX; 8.
SQ SEQUENCE 176 AA; 18434 MW; E84E78E58D435547 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 176;
Best Local Similarity 100.0%; Pred. No. 2.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
Db 50 PPPHG 54

RESULT 35
Q9C4Z8 PRELIMINARY; PRT; 176 AA.
AC Q9C4Z8;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein F27M3.5 (Hypothetical protein F5M6.24).
GN Name=F27M3.5; Synonyms=F5M6.24;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
```


OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE-Testis;
 RL Ansoerge W., Wirkner U., Meves H.W., Gassenhuber J., Wiemann S.;
 RA Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL634172; CAD38870.1; -
 DR HSSP; P42771; 2A5R.
 DR InterPro; IPR002110; ANK.
 DR Pfam; PF00023; Ank; 2.
 DR PRINTS; PR01415; ANKYRIN.
 DR SMART; SM00248; ANK; 2.
 DR PROSITE; PS50088; ANK REPEAT; 1.
 DR PROSITE; PS50297; ANK REP REGION; 1.
 KW ANK repeat; Hypothetical protein.
 FT NON_TER 1
 SQ SEQUENCE 201 AA; 21709 MW; 1E26E07B168873E5 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 201;
 Best Local Similarity 100.0%; Pred. No. 2.8e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
 Db 188 PPPHG 192

RESULT 40
 Q67IV2 PRELIMINARY; PRT; 203 AA.
 AC Q67IV2;
 DT 25-OCT-2004 (TReMBLrel. 28, Created)
 DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
 DT 25-OCT-2004 (TReMBLrel. 28, Last annotation update)
 DE Hypothetical protein OSUNBa0048L03.37.
 GN Name=OSUNBa0048L03.37;
 OS Oryza sativa (japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OX NCBI_TaxID=39947;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sasaki T., Matsumoto T., Fujisawa M.;
 RT "Oryza sativa nipponbare (GA3) genomic DNA, chromosome 6, BAC
 RT clone:OSJNBa0048I03."
 RL Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AP007205; BAD38589.1; -
 KW Hypothetical protein.
 SQ SEQUENCE 203 AA; 21805 MW; 204D6ED4120FD35 CRC64;

Query Match 100.0%; Score 35; DB 2; Length 203;
 Best Local Similarity 100.0%; Pred. No. 2.9e+02;
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PPPHG 5
 Db 9 PPPHG 13

Search completed: June 15, 2005, 14:21:29
 Job time : 85.5 secs